cop 2

OCT 2 1928

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

SEPTEMBER, 1928

Journal of Educational Research

CONTENTS

		Page
A Comparison of the Lecture-Demonstration, Group Laboratory Experimentation, and Individual Laboratory Experimentation Methods of Teaching High School Biology	*Palmer O. Johnson*	103
Minor Studies on Objective Examination Methods—The Negative Suggestion Effect of True-False Tests	*Hazel M. Roberts and G. M. Ruch*	112
A Revision of the Chapman-Sims Socio-Economic Scale	*J. D. Heilman*	117
Scholastic Accomplishment in the Junior High School	*F. C. Landsittel*	127
Mental Ability with Reference to Selection and Retention of College Students	*F. P. OBrien*	136
Waste in Professional Education	*Richard E. Hyde*	144
Resemblance in the Handwriting of Twins and Siblings	*Emily Kramer and Charles E. Lauterbach*	149
Editorials		153
Reviews		155
Research Abstracts and Bibliographies		163
News and Communications		175

Copyright 1928, by Public School Publishing Co.

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879

JOURNAL OF EDUCATIONAL RESEARCH

Published in the interest of better schools and devoted to the development and to the application of tested knowledge of the processes of learning, instruction, and administration. This journal was established to support educational research workers in all phases of the great problems of teaching. Started in 1920 under the editorship of Dr. B. R. Buckingham and associates among the members of an association of directors of educational research, and accepted as the official organ of the Educational Research Association, its purpose has at no time been confined to a narrow line of specialized research but has been concerned also with professional applications for all workers in education.

EDITORIAL BOARD

E. J. ASHBAUGH
Ohio State University

A. S. BARR, Chairman
University of Wisconsin

EARL HUDELSON
University of Minnesota

ASSOCIATE EDITORS

B. R. BUCKINGHAM, *Harvard University*

W. W. CHARTERS, *Director of Bureau of Educational Research, Ohio State University*

M. E. HAGGERTY, *Dean, College of Education, University of Minnesota*

V. A. C. HENMON, *Professor of Psychology, University of Wisconsin*

MURRAY DALMAN, *Educational Consultant, Perkins, Chatten and Hammond, Architects, Chicago, Illinois*

VIRGIL E. DICKSON, *University of California, Berkeley*

CHARLES FORDYCE, *Director, Bureau of Educational Measurements, University of Nebraska*

WILLIAM S. GRAY, *Dean, School of Education, University of Chicago*

KARL J. HOLZINGER, *School of Education, University of Chicago*

ERNEST HORN, *College of Education, University of Iowa*

ARTHUR W. KALLOM, *First Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston*

WILLIAM A. McCALL, *Associate Professor of Education, Teachers College, Columbia University*

GEORGE D. STRAYER, *Teachers College, Columbia University*

W. L. UHL, *Dean, College of Education, University of Washington*

GUY M. WHIPPLE, *Secretary, National Society for the Study of Education*

F. J. KELLY, *President, University of Idaho*

EUGENE NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City*

W. J. OSBURN, *College of Education, Ohio State University*

SHELTON PHELPS, *Director of Instruction, George Peabody College for Teachers*

J. B. SEARS, *Stanford University, California*

W. W. THEISEN, *Assistant Superintendent of Schools, Milwaukee, Wisconsin*

C. W. WASHBURNE, *Superintendent of Schools, Winnetka, Illinois*

CLIFFORD WOODY, *Director, Bureau of Tests and Measurements, University of Michigan*

THE JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Company. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

Subscriptions and all communications about reprints or other business matters should be addressed to the PUBLIC SCHOOL PUBLISHING COMPANY, at Bloomington, Illinois.

Address all communications on editorial matters to A. S. Barr, Department of Education, University of Wisconsin, Madison, Wisconsin.

B. R. Buckingham Takes Up the Next Task

A personality standing each day on the results of that day's work leaves all titles or degrees attained yesterday as passed by in today's progress. It is different with the work itself. That remains always the basis for tomorrow's job.

The plan and purpose which B. R. Buckingham established for the *Journal of Educational Research* may have been all in a day's work for him nine years ago. Its results in that time, however, are found in the present nation-wide recognition of the great value of the entire educational research movement.

The present wide use of accurate means for getting exact facts on pupil difficulties or on their handicaps which limit school success, and for determining their attainments under specific conditions have changed the requirements of service in all lines of educational activity. Undoubtedly the greatest need for school progress today is a better organization of the necessary textbooks.

With his usual insight and his constant devotion to the task next to be done, B. R. Buckingham accepts the call for editorial work in this field of immediate need. The admiration and good wishes of all of his friends go with him.

RECEIPTS AND EXPENDITURES FOR THE JOURNAL OF EDUCATIONAL RESEARCH FROM JULY 1, 1919, TO JULY 1, 1928

Year	Receipts—Subscriptions and Advertisements (1)	Costs—Printing (2)	Paid to Editor (3)	Paid to Ed. Assistant (4)	Expense—Publishing Office and Subscription Agency (5)
1919	$ 579.05				$ 705.00
1920	6515.55	$7231.07	$ 200.00		2171.85
1921	8409.59	8453.77	2100.00		2803.20
1922	13644.05	7190.74	2000.00		4548.00
1923	9898.95	7562.32	2000.00	$125.00	3299.95
1924	9494.07	5922.12	1200.00	400.00	3164.70
1925	9054.51	5327.04		450.00	3018.00
1926	10397.81	6476.64		550.00	3466.00
1927	9226.19	5098.19		450.00	3075.00
1928	3525.15	4328.80		350.00	1175.00
Totals	$86744.92	$57590.69	$7500.00	$2325.00	$27426.70

Total costs and expenses (Columns 2, 3, 4, 5)..... $94842.39
Total receipts (Column 1)..... 80744.92

Excess of costs over receipts for nine years $14097.47
(Average $1566.38 per year.)

Photograph by Bachrach

JOURNAL *of* EDUCATIONAL RESEARCH

Volume XVIII SEPTEMBER 1928 Number 2

A COMPARISON OF THE LECTURE-DEMONSTRATION, GROUP LABORATORY EXPERIMENTATION, AND INDIVIDUAL LABORATORY EXPERIMENTATION METHODS OF TEACHING HIGH SCHOOL BIOLOGY

PALMER O. JOHNSON
University of Minnesota

VARIOUS methods of teaching science have had from time to time ardent supporters and inflexible opponents. It is only recently that attempts have been made to evaluate these claims on an objective basis. Hitherto the prevailing criteria have been teachers' judgments and *a priori* reasoning.

How to teach is a supplement of what to teach. Traditional procedure has been to place chief emphasis upon subject matter; but the increasing concern that is being given to method inspires the hope that it may eventually receive the attention that it merits. For the function of method is the biggest problem directly confronting the classroom teacher. Administrators, school boards, and sometimes even the laity (particularly if they have vested interests) presume to know what should be taught; but the classroom teacher is supposed to know how to teach it. That there may be best ways peculiar to the aims and content prescribed is often given no consideration. Just what technique to employ has been largely determined by the type of training to which the teacher has been subjected. In science this has generally been a transplanting of college methods to the high school situation. Subject matter itself is essentially an abridgment of the college course. Many high school science teachers have been imbued with the attitude that contact with professional literature is supererogative. If they do read or hear of experiments in methods of teaching science the conclusion is often drawn that these methods may be applicable in experimental schools but would not be feasible under ordinary school conditions and with common sense teachers.

If progress in methods is to be made there must be substituted for this state of conditioned emotional response the scientific attitude of the open mind and a willingness to try out different procedures. Progress need

not wait upon laboratory schools. Much valuable information will accrue if teachers in typical schools will test various methods of which they have heard or which may occur to them. The following experiment was an attempt in just this direction; namely, to try out under ordinary school conditions the relative efficacy of three different methods of laboratory instruction in high school biology: the lecture-demonstration, the group experimentation, and the individual experimentation methods. The high school in which the study was made is located in Quincy, Illinois. The enrollment is about eight hundred, which is considerably larger than that of the average high school; but except for more sections per class the conditions pertinent to this study were typical. The fact that there were three sections of biology made it possible to try out three teaching procedures by the rotation plan.

Educational results are interpreted variously. Many theoretical outcomes, as yet incapable of exact measurement, are claimed for every method. Champions of the individual laboratory technique in particular maintain that it results in a number of associate and concomitant outcomes peculiar to that method. Some of their claims are manipulatory skill or laboratory technique, clearer ideas resulting from direct contact with materials, training in problem solving resulting in the general habit of scientific thinking, personal thrills from the mysteries of nature, stimulation of the imagination and development of power of observation, independent thought and action, initiative, judgment, and various other faculties.

All of these are highly desirable; but this study was limited to accomplishments which are capable of quantitative measurement. It consisted of two series of twenty-four experiments, one series being conducted during the first semester, the other during the second. The Terman Group Intelligence Test was given to each section, two tests to each at an interval of two months, and the average used as the basis for mating. In the first series eleven matched students were obtainable in each of the three sections of biology. By the rotation plan each section performed eight experiments by each of the demonstration, individual, and group methods.

In the lecture-demonstration method the instructor performed the experiment, the same directions being followed as were given to the other sections. The order of procedure was (1) statement of the problem; (2) performance according to the method prescribed; (3) making the observations; and (4) formulation of conclusions.

In the group method the class was divided into fours or fives. Each group performed the experiment according to the prescribed method and recorded its observations and conclusions. Freedom of discussion was permitted within the group.

By the individual method each student performed the experiment according to the prescribed method, recorded his observations, and drew his conclusions. No communication was permitted.

The instructor acted mainly as a supervisor in the last two sections, leaving the initiative to the pupils but seeing that the conditions prescribed in the method were rigidly adhered to. Permanent records of the results were required of the members of all sections. Collateral readings were constant for all pupils and were confined to the classroom. A double period of forty minutes each, five times per week was devoted to each section.

The three sections participating in the second series of experiments were made up of sophomores and juniors to the number of 17, 16, and 17, respectively. Their mean intelligence test scores were 137.1 ± 16.8 142.5 ± 24.1, and 144.9 ± 16.2; their mean achievement scores for the first semester's work were 72.7 ± 4.8 (A.D.), 74.1 ± 5.6, and 76.3 ± 5.5.

The examinations were given in some cases immediately following an experiment and sometimes after several experiments, depending upon their nature, but the time was constant for all sections. The tests were all objective in character, consisting of true-false, matching, best answer, completion, and the introduction of diagrams and drawings calling for the identification of apparatus, organisms, and structures. The number of points in the tests ranged from fifty to two hundred. The scores were reduced to a percentage basis. Tests to measure retentivity were administered at various intervals.

The results of the immediate recall tests on the first series of experiments are shown in Table I. The demonstration method excelled the group and individual methods in two of the three sets of experiments;

TABLE I

SUMMARY OF RESULTS OF IMMEDIATE RECALL TESTS, INITIAL EXPERIMENT

EXPERIMENT	SECTION	DEMONSTRATION	SECTION	GROUP	SECTION	INDIVIDUAL
1–8	3	48.4	1	54.4	2	55.1
9–16	2	53.4	3	48.1	1	44.8
17–24	1	53.9	2	50.6	3	46.7
Totals		155.7		153.1		146.6

the individual surpassed the other two methods in one of the three sets. The total mean scores for the series of twenty-four experiments were 155.7, 153.1, and 146.6.[1]

Table II gives the results of the recall tests on the first series of experiments, given one month after the immediate tests. The total mean scores are 126.2, 127.2, and 124.2. The loss in percent of retention is also given, this being a more comparable basis inasmuch as the scores on the immediate tests varied. The mean losses are 19.0, 16.5, and 15.3.

TABLE II

SUMMARY OF THE RESULTS OF DELAYED RECALL TESTS, INITIAL EXPERIMENT (ONE MONTH AFTER IMMEDIATE RECALL TESTS)

	Demonstration			Group			Individual		
Experiment	Section	Score	Percent Loss	Section	Score	Percent Loss	Section	Score	Percent Loss
1— 8	3	37.8	21.9	1	38.8	18.7	2	46.4	15.8
9—16	2	42.9	19.5	3	43.9	8.9	1	35.2	21.4
17—24	1	45.5	15.6	2	44.6	11.8	3	42.6	8.8
Totals		126.2	57.0		127.2	49.4		124.2	46.0
Means		42.1	19.0		42.4	16.5		41.4	15.3

Table III gives a summary of the results of the initial series of experiments on the immediate recall tests by comparing sections and methods. The mean of combined scores for sections 1, 2, and 3 are 51.5 ± 15.7, 52.8 ± 14.9, and 47.6 ± 14.1. The mean combined scores for the various methods are 51.5 ± 13.9, 51.1 ± 14.1, and 48.8 ± 13.9.

TABLE III

COMPARISON BY SECTIONS AND METHODS OF THE RESULTS OF THE IMMEDIATE RECALL TESTS IN THE INITIAL EXPERIMENT

	Section 1		Section 2		Section 3		Combined Scores	
Method	Mean	S.D. dis.	Mean	S.D. dis.	Mean	S.D. dis.	Mean	S.D. dis.
Demonstration	53.9		53.4		48.4		51.5	± 13.9
Group	54.4		50.6		48.1		51.1	± 14.1
Individual	44.8		55.1		46.7		48.8	± 13.9
Mean of Combined Scores	51.1	± 15.7	52.8	± 14.9	47.6	± 14.1		

A summary of the results on the immediate tests of the second series of experiments is given in Table VI. It will be seen that the demonstration method surpassed the other two in all three sets of experiments. The group excelled the individual method in two of the sets. The total scores are 198.58, 188.02, and 185.17.

[1] For the demonstration, group, and individual methods, respectively. This order will be followed throughout.

TABLE IV

SIGNIFICANCE OF OBTAINED DIFFERENCES IMMEDIATE RECALL TESTS, INITIAL EXPERIMENT

METHOD	DIFFERENCE	S.D. DIFFERENCE	DIFFERENCE / S.D. DIFFERENCE
Demonstration–Individual	2.7	3.58	.75
Demonstration–Group	.4	3.6	.11
Individual–Group	2.3	3.6	.64

TABLE V

SIGNIFICANCE OF OBTAINED DIFFERENCES BETWEEN SECTIONS IMMEDIATE RECALL TESTS, INITIAL EXPERIMENT

SECTION	DIFFERENCE	S.D. DIFFERENCE	DIFFERENCE / S.D. DIFFERENCE
Section 1–Section 2	1.7	3.86	.44
Section 1–Section 3	3.5	3.78	.92
Section 2–Section 3	5.2	3.67	1.40

TABLE VI

SUMMARY OF RESULTS OF IMMEDIATE RECALL TESTS, SECOND EXPERIMENT

EXPERIMENT	SECTION	DEMONSTRATION	SECTION	GROUP	SECTION	INDIVIDUAL
1– 8	1	62.32	2	59.05	3	60.99
9–16	2	65.84	3	63.71	1	59.96
17–24	3	70.42	1	65.26	2	64.22
Totals		198.58		188.02		185.17

TABLE VII

SUMMARY OF RESULTS OF DELAYED RECALL TESTS, SECOND EXPERIMENT

		DEMONSTRATION			GROUP			INDIVIDUAL		
EXPERIMENT	ELAPSED TIME	Sec.	Score	Percent Loss	Sec.	Score	Percent Loss	Sec.	Score	Percent Loss
1– 8	4 weeks	1	55.72	10.59	2	54.31	8.03	3	56.62	7.16
9–16	3 weeks	2	58.82	10.66	3	57.77	9.32	1	50.34	16.04
17–24	2 weeks	3	65.79	6.57	1	59.69	8.54	2	59.47	7.39
Totals			180.33	27.82		171.77	25.89		166.43	30.59
Means			60.11	9.27		52.26	8.63		55.48	10.2

Table VII presents a summary of the results of the recall tests in the second series of experiments. On the basis of percent of information retained, each method occupied a first, second, and third rank. The section highest in intelligence and achievement rating stood first in all three sets of experiments and the section second in intelligence and achievement stood second in all three sets. The mean losses in percent of retention are 9.27, 8.63, and 10.2.

Table VIII gives a summary of the immediate recall test results of the second series of experiments by comparing sections and methods. Total scores for sections 1, 2, and 3 are 187.54 ± 24.4, 189.11 ± 25.9, and 195.12 ± 28.8. The mean combined scores for the various methods are 66.3 ± 10.8, 62.58 ± 9.5, and 61.76 ± 10.6.

TABLE VIII

COMPARISON BY SECTIONS AND METHODS OF THE RESULTS OF IMMEDIATE RECALL TESTS IN THE SECOND EXPERIMENT

	Section 1		Section 2		Section 3		Combined Scores	
Method	Mean	S.D. dis.	Mean	S.D. dis.	Mean	S.D. dis.	Mean	S.D. dis.
Demonstration	62.32	± 8.13	65.84	± 8.84	70.42	± 9.84	66.30	± 10.78
Group	65.26	± 9.60	59.05	± 8.26	63.71	± 9.68	62.58	± 9.52
Individual	59.96	± 10.80	64.22	± 9.24	60.99	± 11.18	61.76	± 10.60
Totals	187.54	± 24.4	189.11	± 25.9	195.12	± 28.8		

Table IX presents a summary of the results of the second series of experiments in the delayed recall tests by comparing sections and methods. Total scores for sections 1, 2, and 3 are, respectively, 165.75 ± 28.05, 172.6 ± 28.13, and 180.18 ± 27.25. The mean combined scores for the three methods are 60.16 ± 10.94, 57.4 ± 10.02, and 55.91 ± 10.91.

TABLE IX

COMPARISON BY SECTIONS AND METHODS OF RESULTS OF DELAYED RECALL TESTS IN SECOND EXPERIMENT

	Section 1		Section 2		Section 3		Combined Scores	
Method	Mean	S.D. dis.	Mean	S.D. dis.	Mean	S.D. dis.	Mean	S.D. dis.
Demonstration	55.72		58.82		65.79		60.16	± 10.94
Group	59.69		54.31		57.77		57.40	± 10.02
Individual	50.34		59.47		56.62		55.91	± 10.91
Totals	165.75	± 28.05	172.60	± 28.13	180.18	± 27.25		

In order to determine the reliability of the differences obtained between the various methods, due to a limited number of samples, it is necessary to measure the probable divergence of the true difference from the obtained difference. The unreliability of a difference between A and B equals the square root of the sum of the square of the unreliability of A and the square of the unreliability of B.[2] The data for the significance of obtained differences are given in Tables IV, V, X, XI, XII, and XIII.

If the criterion of a significant difference is taken such that $\frac{\text{obtained difference}}{\text{S.D. difference}} = 3$, it is observed that no obtained difference will be statistically significant. The nearest approach is that between the demonstration and individual methods as shown in Table X, where $\frac{\text{difference}}{\text{S.D. difference}} = 2.12$. From tables of values of the normal probability integral[3] it is found that between the average, 0, and $\pm 2.12 \left(\frac{\text{difference}}{\text{S.D. difference}}\right.$

[2] Thorndike, E. L. *Mental and Social Measurements*, pp. 188-93.
[3] *Ibid.*, p. 198.

for the difference between the demonstration and individual methods) there are 48.3 percent of the cases. The facts are identical for -2.12. This means that the true difference will not differ from the obtained difference, 4.54, either above or below, by more than 2.38, in more than 34 cases out of a thousand.

TABLE X

SIGNIFICANCE OF OBTAINED DIFFERENCES, IMMEDIATE RECALL TESTS, SECOND EXPERIMENT

METHOD	DIFFERENCE	S.D. DIFFERENCE	DIFFERENCE / S.D. DIFFERENCE
Demonstration–Individual	4.54	2.14	2.12
Demonstration–Group	3.72	2.03	1.80
Individual–Group	.82	2.01	.408

TABLE XI

SIGNIFICANCE OF OBTAINED DIFFERENCES BETWEEN SECTIONS IMMEDIATE RECALL TESTS, SECOND EXPERIMENT

SECTION	DIFFERENCE	S.D. DIFFERENCE	DIFFERENCE / S.D. DIFFERENCE
Section 1–Section 2	1.57	5.06	.3
Section 1–Section 3	7.58	5.28	1.4
Section 2–Section 3	6.01	5.5	1.1

TABLE XII

SIGNIFICANCE OF OBTAINED DIFFERENCES BETWEEN METHODS DELAYED RECALL TESTS, SECOND EXPERIMENT

METHOD	DIFFERENCE	S.D. DIFFERENCE	DIFFERENCE / S.D. DIFFERENCE
Demonstration–Individual	4.25	2.19	1.9
Demonstration–Group	2.76	2.12	1.3
Individual–Group	1.49	2.13	.7

TABLE XIII

SIGNIFICANCE OF OBTAINED DIFFERENCES BETWEEN SECTIONS DELAYED RECALL TESTS, SECOND EXPERIMENT

SECTION	DIFFERENCE	S.D. DIFFERENCE	DIFFERENCE / S.D. DIFFERENCE
Section 1–Section 2	6.85	5.7	1.2
Section 1–Section 3	14.43	5.5	2.6
Section 2–Section 3	7.58	5.6	1.4

Interpretations.—Although the demonstration method outranked the others in five of the six sets of experiments, it cannot be conclusively stated that it is superior, since the obtained differences were not found to be statistically significant. However, it is apparent that the demonstration method of laboratory instruction may be expected to yield, if not larger, at least equal returns in primary learning when compared

with the group or individual method. This becomes significant in view of the saving in laboratory equipment.

There appears to be no significant difference between results obtained from the group and individual methods.

With respect to the permanence of accomplishment, there is a suggestion that native retentivity may be a more influential factor than is the teaching method, with probably a slight tendency for the demonstration to excel the group and individual methods.

In the light of the standard deviations in the scores attained by the various methods, there appears to be a tendency toward a wider deviation by the individual method than by either of the others. The individual method, therefore, seems to provide greater opportunity for the exercise of individual differences.

Implications.—The outstanding result of this study is that the three particular methods of teaching seem to differ very little in their influence upon pupil learning and still less upon pupil retention. No one procedure excelled in all of the experiments. It may well be that each method has its place in science teaching. If so, the proper function of each should be determined. A vote of the pupils suggested that either the method should be varied or different methods should be used with different pupils, for there were twenty-five votes for the group method, twelve for the demonstration, and eleven for the individual.

It may be, on the other hand, that the failure to reveal a clear superiority for any one method was due to limitations in the experimental technique employed. A larger number of cases would certainly have been desirable. With limited numbers there may be a better single basis of pairing than intelligence test scores.[4] Possibly achievement was not measured with sufficient accuracy to reveal tendencies. Perhaps a measure of improvement between initial and final tests would have produced different results. Possibly scaled test results would better have lent themselves to statistical treatment than do raw scores. Industry, attention, interest in science, previous training, and variation in the appeal of teaching methods to pupils may all be influential factors. Moreover, the same methods might produce quite different results under another teacher. It is entirely reasonable to assume that they would; consequently the duty of each teacher is to determine experimentally his own best methods of teaching. It is neither necessary nor advisable for

[4] The correlation between intelligence test scores and achievement was .49 ± .12.

him to wait upon laboratory schools to determine the relative values of methods, for the experimental conditions may be quite different and one teacher's forte may not be another's at all.

The seeming difference of results in the present study may, however, mask their true significance. Since the demonstration method, which at least held its own as a means to immediate learning, makes possible a considerable saving in apparatus and a very decided saving of pupil time, it may be able to more than offset any advantages that either of the other methods can offer. In these respects the individual method, which now prevails, ranks lowest. If it is to persist it should be required to justify itself on other grounds. It has not assuredly done so in this study.

MINOR STUDIES ON OBJECTIVE EXAMINATION METHODS

I. The Negative Suggestion Effect of True-False Tests

HAZEL M. ROBERTS AND G. M. RUCH
University of California

Historical and Introductory.—The belief in the supposed negative suggestion phenomenon arising from the answering of the false items of a true-false test rests entirely upon *a priori* grounds as far as the writers are aware. Prior to the present investigation, the only experimental study of this question seems to be that of Remmers and Remmers.[1]

That false statements, if presented to pupils, *might* leave their residues of misinformation seems plausible enough in theory. The real problem is to ascertain the facts. While we are still considering theoretical issues, the authors wish to call attention to certain related considerations.

1. If there is a negative suggestion effect in false statements, it is equally likely that there are positive suggestion (or teaching) effects in true statements. If so, the pragmatic problem may be that of ascertaining the *net* effect of these two phenomena.

2. The negative suggestion effects, to be serious, must be *permanent*. For such false teaching to persist a day, a week, a month, or even a year, might be quite without significance in the long run.

3. A careful distinction must be made between two quite different situations: (a) the answering of true-false statements where instructions are given to the effect that "approximately half of the following statements are true and about half are false," and (b) the studying of a printed page in a text containing errors when such learning is done in "good faith." The difference resides in the *mental set*. In the first case there is a critical, challenging, attitude ("Bewusstseinslage"); in the second situation there is uncritical acceptance of error due to the authority and sanctity of the printed page.

4. It seems unlikely that false ideas would be "fixed" in mind when the pupil possesses positive knowledge of the fact; the negative suggestion effect, if it exists, will be largely confined to those test items upon which the pupil's knowledge is hazy or entirely absent. *In such a case negative teaching represents the very unlikely psychological situation of learning through a single impression.*

[1] Remmers, H. H. and E. M. "The Negative Suggestion Effect of True-False Examination Questions," *Journal of Educational Psychology*, XVII (1926), 52-56.

5. There is sound logic in holding that teaching should, in part, consist in distinguishing between truth and error since life presents these in scrambled fashion. Moral education through "running away from evil without a battle" is now generally held to have proved a failure.

Remmers and Remmers gave an unfamiliar reading passage to two groups of 68 college students each, the groups having been equated for intelligence. Group I took immediately 121 true-false items based upon the passage. Group II answered 121 recall items over the same material. About four weeks later, and without warning, these two groups were retested with the tests reversed. Conclusions were based upon the sizes of the mean scores at the two testings upon the logic that, if upon the second testing, the mean score of Group I (which had previously taken the true-false test) was as high as or higher than that of Group II (which had not taken the true-false test previously), the true-false test had left no permanent residue of misinformation. Such, indeed, was the experimental finding. The slight differences found indicated that the true-false test facilitated delayed recall—a *positive* suggestion effect in the terminology of the present writers.

The Authors' Experiments.—Two types of experiments were carried out. The first, not reported in detail here, used the procedure of the Remmers except that high-school pupils in botany and chemistry were the subjects. The results showed no negative suggestion effect in botany and a slight one in chemistry although the odds for the latter were only four to one that the differences were statistically significant, indicating a difference far too slight for acceptance. The only conclusion that can be drawn is that the negative suggestion effect is too small for practical consideration, if existent at all. These findings are in substantial agreement with those of Remmers and Remmers.

In the second and major experiments the procedure was as follows:

1. A completion test was given first.

2. A true-false test covering the "same" items was given immediately after the completion test.

3. From one day to one month later, the same completion test was repeated.

4. The two completion tests were compared, item by item, to see what changes were made (presumably as a result of the intervening true-false test). Detailed tabulations were made in order to show whether the pupils accepted the identical answer (of the true-false test), whether

right or wrong, in the second completion test, or whether some other change (not suggested by the true-false) was made.

5. As a general check, second examinations of different test items were given exactly as before except that no true-false tests at all were administered, i.e., the testing consisted of an initial completion test followed after an interval with a repetition of the same test. This plan was used as a rough measure of general practice effects and as a check on the looking-up of answers between tests. (See Part III of Table I.)

Table I shows the general results.

TABLE I

SHOWING THE DIFFERENCES BETWEEN THE SCORES ON THE INITIAL COMPLETION TESTS AND THE SAME TESTS REPEATED AT VARYING INTERVALS WITH TRUE-FALSE TESTS COVERING THE "SAME ITEMS" INTERVENING

Section	Interval	Difference in Averages	S.D. of Difference	Difference / S.D. of Diff.	Nos.
I. *Short Interval*					
Biology A	1	8.2	0.6	14.2*	21
Chemistry IA	3	4.1	0.8	5.0	20
Chemistry SB	3	1.8	0.4	5.0	26
Chemistry SA[1]	4	2.1	0.5	4.1	24
Physics B	5	3.1	0.4	7.3	24
				Total	115
II. *Long Interval*					
Biology B	15	0.4	1.1	0.4	17
Chemistry IB	20	2.2	0.7	3.0	21
Chemistry NA	12	2.5	0.7	3.7	19
Chemistry NB	31	3.0	0.7	4.2	20
Physics C	33	0.8	0.8	1.0	22
Physics A	34	—0.4	0.7	—0.5	21
				Total	120
III. *Long Interval, no Intervening True-False Test*					
Botany A, Exam. 2	14	1.9	0.8	2.3	33
Botany B, Exam. 1	15	3.2	1.1	2.9	28
Chemistry A[1]	14	0.0	0.6	0.1	22
Chemistry A[3]	14	—1.3	0.8	—1.6	21
Chemistry B	14	1.6	0.6	2.4	24
				Total	128

* The ratios in this column cannot be obtained exactly from the values in preceding columns since the computations were done with two-places of decimals and subsequently rounded off to one.

Discussion of Results of Table I.—Since there is no valid way of pooling all the sections into a single measure, Table I must be studied for the general sweep of the evidence.

1. With two exceptions (Sections *Physics A* and *Chemistry A*[2], *Examination C*), the second completion test gave higher mean scores, thus indicating a gain from some cause or causes, probably (a) practice effects and looking-up of answers or (b) *net* positive suggestion effects.

2. If we accept differences of more than three times their standard errors as being statistically significant, there is a net positive effect in all five of the short interval comparisons. Also three out of six of the longer interval comparisons show positive suggestion effects.

3. When no true-false test intervened, out of five comparisons, not one showed a statistically significant gain although three of the five showed gains of more than two times the sigma of such gains. This would indicate that practice effects from repeating the completion test will not *alone* explain the gains of Table I. There is probably in addition a net positive suggestion effect of true-false tests.

4. The net positive suggestion effect is probably transitory since the differences after a short interval are from four to fourteen times their standard errors but those after a long interval (13 to 34 days) never exceed four times their standard errors and in half the cases are not even statistically significant. It may reasonably be supposed that, after intervals of a few months or a year, the net positive suggestion effect would be entirely lost. This is reasonable in view of the fact that but *two* experiences with the item in completion form and *one* experience with it in true-false form represented the learning situation.

5. The evidence for any net negative suggestion effect is quite lacking unless the rather unreasonable view is taken (in spite of the evidence of Part III of Table I) that practice effects have masked the results.

6. It should be noted that we are discussing *net* results, either positive or negative, as far as we are concerned with the phenomenon of suggestion. We shall see later that negative effects do occur in many items, but that these are more than outweighed by positive effects, according to our results.

Results From the Tracing of Answers to Individual Test Items Through the Three Testings.—It has been mentioned that each item was followed through the three tests, the initial completion, the true-false, and the second completion after varying intervals. The detailed results would require many pages of tabulations. Instead we shall confine the present discussion to a summary of Miss Robert's findings.[2] The principal results are:

[2] The full text and tabulations are on file in the Lange Library, University of California, Berkeley, California, under the title: Roberts, Hazel M. "Have True-False Examinations a Measurable Negative Suggestion Effect." (M. A. Thesis, 1927).

1. About 1 right answer (initial completion test) in 20 was changed to the identical wrong suggested by the intervening true-false test. At least one-third of these changes seemed due to chance.

2. About 1 wrong answer (initial completion test) in 7 was changed to the identical wrong suggested by the intervening true-false test. About one-third of these may be held to be due to chance.

3. About 1 omitted answer (initial completion test) in 9 was changed to the identical wrong answer of the intervening true-false test. Perhaps all of these may be charged to the influence of the false statements.

4. The three foregoing statements were markedly less true of the retestings after the longer intervals. This would indicate that negative suggestion effects were transitory in the main.

5. As suggested by Table I, the changes from rights or omissions (initial completion tests) to the identical true of the intervening true-false test are more numerous than are the changes to identical wrongs.

GENERAL CONCLUSIONS

1. The negative suggestion effect exists, at least for short intervals.

2. It is very much smaller than many critics, reasoning *a priori*, believe.

3. The *net* suggestion effect of a true-false test appears to be *positive*, not negative.

4. The existence of a net positive suggestion effect, if further investigations agree, does not eliminate the entire force of the objection to true-false tests on this score since certain misinformations may be fixed in spite of the net beneficial instructional effects.

5. Negative suggestion effects, positive suggestion effects, and the *net* positive suggestion effects appear to be rather transitory as intervals of two weeks to a month between testings gave markedly smaller gains than did intervals of from one to five days. It seems quite within reason that an interval of one or two school semesters might wipe out such effects completely.

A REVISION OF THE CHAPMAN-SIMS SOCIO-ECONOMIC SCALE[1]

J. D. HEILMAN
Colorado State Teachers College

IN connection with an investigation made to determine the relative influence upon educational achievement of some hereditary and environmental factors,[2] it was felt desirable to have some objective measure of the cultural, social, and economic background of the children who were selected as the subjects of the investigation. It was felt desirable, not only for the purpose of determining the influence of the socio-economic status of the child's home upon his educational achievement, but also to avoid attributing too large a share of influence upon educational age to the other factors—school training and mental age.

The Chapman-Sims Scale appeared to be the most objective and satisfactory measure available. For several reasons, the scale was not used in its original form. As its construction was based upon selected homes of New Haven, an Eastern city, the items of the scale might receive different relative weights if based upon homes of a city of the Rocky Mountain region. As a matter of fact, it was found that one of the items, if used as prescribed by the Chapman-Sims Scale, lost much of its value in discriminating the poorer from the better homes. Moreover, the sampling of the homes used in the Chapman-Sims Scale was thought to be inadequate, as only those homes were involved whose children attended the New Haven academic high school. Finally, the important items of "the father's occupation" and "the number of rooms occupied by the family" were omitted from the scale, because the question on the former was misunderstood by the children, and because the latter could not be used on account of the apartment house situation. These were included in our revised scale.

The data for the revised scale were, with few exceptions, collected from the homes of all of the ten-year-old children attending 48 public

[1] Chapman, J. C. and Sims, V. M. "The Quantitative Measurement of Certain Aspects of Socio-economic Status." *Journal of Educational Psychology*, Vol. XVI (September, 1925). (Editor's Note: It should be noted that Mr. Sims has himself just recently, 1927, issued through the Public School Publishing Company a revised Form C of the "Sims Score Card for Socio-Economic Status."—G.M.W.)

[2] Heilman, J. D. *Twenty-Seventh Yearbook of the National Society for the Study of Education*, 1928, Part II, Ch. II.

schools in Denver, Colorado. The exceptions were the homes of those children whose parentage was foreign, whose records were incomplete, and who had attended other than the Denver schools. The 48 schools were selected on the basis of the average socio-economic status of the school communities. This may have been a disadvantage, but it is probable that both the very poorest and the very best homes were represented in these communities. At least, there were a few homes which received the lowest possible score on the scale and a relatively large number of homes which received the highest possible score.

Data on socio-economic status were obtained for about 1000 children. This number was reduced to 828, because there was insufficient time for the examiners to obtain the mental ages of 172. To determine the scale values only 688 of the 828 socio-economic status records were used, because the remaining records were incomplete. Absence of the father's occupation, owing to decease or separation, was the chief cause of incomplete records.

QUESTION BLANK FOR PARENTS

NAME OF CHILD..... DATE..... SCHOOL.....

1. Do you have a telephone in your home?.....
2. Did the father of the child attend high school?.....
3. Did the mother of the child attend high school?.....
4. Is English the only language spoken in the home?.....
5. Is your house heated by a furnace?.....
6. Do you have an auto other than a truck?.....
7. Do you have a piano in your home?.....
8. Do you have a phonograph in your home?.....
9. How many daily papers do you take regularly in your home?.....
10. How many magazines do you take regularly in your home?.....
11. About how many books are there in your home?.....
12. Do you live in an apartment house?.....
13. How many rooms, not including bath room, are there in your home?.....

.....

(Name of Parent)

The data were collected from the homes by the above questionnaire, which was accompanied by a letter signed by Charles E. Greene, Director of Measurements of the Denver Schools. The letter urged the parents to coöperate by filling in the questionnaire accurately and completely. It was also designed to allay any fears which the parents might have, by telling them that the results would not be seen by the teachers and that they would not in any way affect the standing of their children, but would be used only in a study made by the National Society for the Study of Education on the importance of regular school attendance.

Item 12 was considered in connection with Item 13. If Item 12 was answered in the affirmative, Item 13 was not used in determining the score. The father's occupation was obtained from the school records. The entire scale, therefore, is composed of thirteen items. These items are to be regarded as indices of the cultural, social, and economic conditions of the home. No claim is made that these thirteen items are the most significant which could have been selected. A better scale could probably be made by obtaining answers to 30 or 40 items, followed by an elimination of those of minor importance. One weakness of the list of questions is that it provides for recording the possession of a phonograph, but not for the possession of a radio. A fairly large number of people stated that they had no phonograph, but did own a radio. Evidently, in some homes the radio is taking the place of the phonograph.

Questions 1 to 8, inclusive, are answered by "yes" or "no," i.e., possession or non-possession. As the other questions could not be answered in this manner, and as it was desired to treat all of the responses uniformly, it became necessary to determine for these others a point that would separate 'possession' from 'non-possession.' It was thought desirable so to select these points that about 65 percent of all the homes would belong to the possession group, but it was found impossible to do this. For example, if one daily paper or more was selected as the distinguishing point, then 98 percent of the homes received a possession score; but if two or more papers were taken as the distinguishing point, then only 33 percent of the homes received a score for possession. The points were therefore arbitrarily chosen at such a place as seemed likely to give the item the highest efficiency possible as a measure of socio-economic status.

The distinguishing point for "papers" was taken at two or more; for "magazines" at three or more; for "rooms" at six or more; and for "occupation" at a point corresponding to a rating of 11.17 or more on the Barr Occupation Scale.[3] This point for occupation gave the possession score to all professional, business, and clerical groups, and the non-possession score to sales clerks and to the skilled, semi-skilled, and common-labor groups.

After the distinguishing points for possession and non-possession had been established, the percentage of homes receiving possession and non-

[3] Barr, F. C. *A Scale for Measuring Mental Ability in Vocations and Some of Its Applications.* (Unpublished Master's Thesis, Leland Stanford Junior University, 1918.) See also Terman, L. M. *et al. Genetic Studies of Genius,* Vol. I., (Stanford University Press, 1925).

possession scores on each item was computed. On the basis of these percentages, and with the aid of the Kelley-Wood Normal Probability Integral Table,[4] possession and non-possession for each item were transmuted into sigma deviations from the arithmetic mean of the entire group. In order to eliminate negative quantities, five units were added to each of the sigma values obtained for a given item. Decimals were then eliminated by multiplying the sigma values by ten. The results of the computations described in this paragraph are given in Table I.

TABLE I

PERCENTAGE OF POSSESSION AND NON-POSSESSION SCORES FOR EACH SOCIO-ECONOMIC INDEX AND THE CORRESPONDING VALUES OF POSSESSION AND NON-POSSESSION IN TERMS OF THE STANDARD DEVIATION FROM THE ARITHMETIC MEAN FOR 688 DENVER HOMES

INDEX	PERCENTAGE		S.D. DEVIATIONS		S.D. DEVIATIONS + 5		S.D. VALUES × 10	
	P	N-P	P	N-P	P	N-P	P	N-P
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Telephone	71	29	.48	—1.18	5.48	3.82	55	38
2. Father	51	49	.78	—0.81	5.78	4.19	58	42
3. Mother	53	47	.75	—0.85	5.75	4.15	58	42
4. English	91	9	.18	—1.80	5.18	3.20	52	32
5. Furnace	63	37	.60	—1.02	5.60	3.98	56	40
6. Automobile	67	33	.54	—1.10	5.54	3.90	55	39
7. Piano	69	31	.51	—1.14	5.51	3.86	55	39
8. Phonograph	64	36	.58	—1.04	5.58	3.96	56	40
9. Papers (2+)	34	66	1.08	—0.56	6.08	4.44	61	40
10. Magazines (3+)	58	42	.67	—0.93	5.67	4.07	57	41
11. Books (125+)	41	59	.95	—0.66	5.95	4.34	60	43
12. Rooms (6+)	53	47	.75	—0.85	5.75	4.15	58	42
13. Occupation	48	52	.82	—0.77	5.82	4.23	58	42

For use in further work on the construction of the scale, a tentative socio-economic status score was computed for each child. This method was to assign to each item one of the values found in the last two columns of Table I and then to find the sum of the thirteen values. If the child received a possession score for an item, the value was found in the second last column; if he received a non-possession score, the value was found in the last column. If, for example, a child possessed the telephone, the item was given a value of 55; if he did not, a value of 38.

The next step in the construction of the scale was to determine, for the values of the last two columns of Table I, weights based upon their relative importance in determining socio-economic status. In order to determine their relative importance, the tentative socio-economic status score was used as the criterion. The scores for the 688 children on each item were in turn correlated with the scores obtained from the sum of the

[4] Kelley, T. L., *Statistical Method* (The Macmillan Company, 1923), p. 373.

remaining items, thus giving thirteen coefficients of correlation. The items yielding the highest coefficients with the criterion were then given the heaviest weighting. A more refined method of determining these weights would have been to compute coefficients of determination, but the rough nature of the scale, together with the enormous labor involved and the lack of knowledge as to the extent of the reciprocal influences of the several items, made such an undertaking impracticable.

For the purpose of computing the coefficients used in weighting the values in the last two columns of Table I, the Pearson formula[5] for the bi-serial r method was employed.

The coefficients obtained for each item or index are given in Table II. The coefficients found by Chapman and Sims for the items common to the two scales are also given.

TABLE II

COEFFICIENTS OF CORRELATION BETWEEN EACH ITEM OF THE SCALE AND THE TOTAL OF THE REMAINING ITEMS

INDEX	BI-SERIAL r	
	Revised Scale	Original Scale
(1)	(2)	(3)
1. Telephone	.77	.54
2. Education of father	.60	.50
3. Education of mother	.69	.67
4. Use of languages	.26	.46
5. Furnace	.75	.48
6. Automobile	.54	.52
7. Piano	.50	.59
8. Phonograph	.39	.36
9. Number of daily papers	.27	.28
10. Number of magazines	.63	.58
11. Number of books	.73	.62
12. Number of rooms	.58	...
13. Occupation of father	.71	...

The degree of correlation between the two series of coefficients given in Table II is not very high, only .59. This may be due in part to the fact that the revised scale contains two items which are not included in the original scale, and the original scale contains two items which do not appear in the revision. Moreover, the sampling of homes for the original scale was very different from the sampling used in the revised scale. Some difference may also be due to the fact that Chapman and Sims used only a sample of 100 cases in computing their coefficients while I used all of the 688 records. At first, I used a random sample of 177 records, but was somewhat dissatified with the results because the distribution did not conform very closely to the expected distribution, having too many cases

[5] Kelley, T. L. *Statistical Method* (The Macmillan Company, 1923), p. 245.

at the upper end. However, the series of co-efficients obtained from the sample correlated with the series based on the entire population to the extent of .98. The lack of correspondence between the series of coefficients for the original and revised scales is marked only for the items "telephone," "furnace," and "piano."

For the revised scale all of the coefficients, excepting those for Items 4, 8, and 9, lie between .50 and .77. All but three of the coefficients are higher than those of the original scale.

The method of weighting on the basis of the size of the coefficients gave more recognition to small differences than did the method employed by Chapman and Sims, although no claim is made that their method is inferior, because the nature of the scale may not require fine distinctions in weighting. The scheme of weighting may be seen from the following tabulation:

Coefficients	Weights
.75 +	1.00
.70 — .74	.85
.65 — .69	.7
.60 — .64	.6
.55 — .59	.5
.50 — .54	.4
.45 — .49	.4
.40 — .44	.3
.35 — .39	.3
.30 — .34	.3
.25 — .29	.2

The actual weighting of each item or index and the final sigma values found for the possession and non-possession of the items are shown in Table III.

TABLE III

The Weighting of Each Item or Socio-Index on the Basis of the Size of its Co-efficient with the Criterion

Index	Original Values		Weights	Original Values × Weights		5 Added to Columns 5 and 6		Final Values ×10	
	P	N-P				P	N-P	P	N-P
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Telephone	.48	−1.18	1.0	.48	−1.18	5.48	3.82	55	38
2. Father	.78	− .81	.6	.468	− .486	5.47	4.51	55	45
3. Mother	.75	− .85	.7	.525	− .595	5.53	4.41	55	44
4. English	.18	−1.80	.2	.036	− .360	5.04	4.64	50	46
5. Furnace	.60	−1.02	1.0	.600	−1.020	5.69	3.98	56	40
6. Automobile	.54	−1.10	.4	.216	− .440	5.22	4.56	52	46
7. Piano	.51	−1.14	.4	.204	− .456	5.20	4.54	52	45
8. Phonograph	.58	−1.04	.3	.174	− .312	5.17	4.69	52	47
9. Papers (2+)	1.08	− .56	.2	.216	− .112	5.22	4.89	52	49
10. Magazines (3+)	.67	− .93	.6	.402	− .558	5.40	4.44	54	44
11. Books (125+)	.95	− .66	.85	.808	− .561	5.81	4.44	58	44
12. Rooms (6+)	.75	− .85	.5	.375	− .425	5.38	4.58	54	46
13. Occupation	.82	− .77	.85	.697	− .655	5.70	4.35	57	44
Total								702	578

The values in the ninth column of Table III are the final values of the scale to be assigned for the possession of an item, and the values in the tenth column those to be assigned for the non-possession of an item. It is not the magnitude of these values which fixes the relative weight of the different items of the scale, but the magnitude of the difference between the possession and non-possession values. In Table IV, the items are arranged in order on the basis of these differences. For the sake of comparison, I am giving the difference values of the original scale for each item.

TABLE IV

ITEMS OF THE SCALE ARRANGED IN THE ORDER OF THEIR RELATIVE WEIGHT IN THE SCALE

ITEM OR INDEX	DIFFERENCE BETWEEN POSSESSION AND NON-POSSESSION VALUES	
	Revised Scale	Original Scale
(1)	(2)	(3)
1. Telephone	17	13
2. Furnace	16	13
3. Number of books	14	17
4. Occupation	13	
5. Education of mother	11	16
6. Education of father	10	13
7. Number of regular magazines	10	16
8. Number of rooms	8	
9. Piano	7	17
10. Automobile	6	13
11. Phonograph	5	10
12. Use of languages	4	13
13. Number of daily papers	3	6
Total	124	147

A comparison of these difference values will show that the items are scaled very differently in the two scales. Which scale is the better has not been finally decided. The revised scale yields the higher reliability coefficient, but to find this coefficient, the items were not grouped in the same way as in the original scale. Chapman and Sims used two methods of grouping. In one, the cultural items were correlated with the economic and in the other a random grouping was employed. The latter yielded a coefficient of .63, or .77 when the Brown formula[6] for computing the coefficient of two equivalent scales was employed. The corresponding coefficients for the former method are .49 and .66, respectively. For the revised scale, an attempt was made to pair a cultural item in one group with a cultural item in the other. The economic items were similarly paired. Moreover, they were so grouped that the total difference values

[6] Brown, W. *The Essentials of Mental Measurement* (Cambridge, England, 1911), p. 102.

of one group equalled the difference values of the other. The two groups are given in Table V. The coefficient obtained for one-half with the other was .77, and for two similar scales as determined by the Brown formula .87. On account of a fairly large number making a perfect score this coefficient may be somewhat higher than it should be.

TABLE V

GROUPING OF ITEMS FOR COMPUTING COEFFICIENT OF RELIABILITY

GROUP I		GROUP II	
Item (1)	Difference Values (2)	Item (3)	Difference Values (4)
Telephone	17	Furnace	16
Education of father	10	Education of mother	11
Number of daily papers	3	Use of languages	4
Occupation	13	Number of rooms and automobile	14
Phonograph	5	Piano	7
Number of books	14	Number of regular magazines	10
Total	62	Total	62

Because the importance or weight of an item is determined by the difference between the possession and non-possession values, the scale can be simplified by giving a zero score for non-possession and the difference-value of an item, as given in Table IV, for possession. On account of the impossibility of securing responses from all homes on every item, it would be well to find the average for all of the scores obtained. In making use of the scale in the investigation for which it was revised, the scores were averaged and multiplied by ten. This reduced the range of the scale from 578-702 to 445-540. The distribution of the scores for the 828 homes involved in the investigation to determine the relative influence upon educational achievement of some hereditary and environmental factors, is given in Table VI. Percentile points are also given.

The distribution table shows that there are too many frequencies at the high end, as 50 cases made a perfect score. The skewed nature of the curve is probably chiefly due to the selection of the homes in the communities of average social standing. Evidently, there were many of the very best families in these communities. Perhaps it would be desirable to make the scale more difficult at the upper end. The results of an extensive application of the scale to unselected homes are needed to discover what changes in the scale are necessary. The scale can undoubtedly be much improved by further attempts at revision.

TABLE VI

FREQUENCY DISTRIBUTION OF SCORES MADE ON SOCIO-ECONOMIC SCALE BY 828 DENVER HOMES SELECTED FROM COMMUNITIES OF AVERAGE SOCIAL AND ECONOMIC STANDING, AND PERCENTILE POINTS

SCALE VALUES (1)	FREQUENCIES (2)		PERCENTILES (3)	POINTS (4)
540	50		10	540
535	49	99	9	536
530	51			
525	55	106	8	528
520	41			
515	50	91	7	519
510	44			
505	54	98	6	511
500	40			
495	54	94	5	502
490	33			
485	44	77	4	493
480	34			
475	38	72	3	482
470	36			
465	46	82	2	471
460	22			
455	40	62	1	459
450	23			
445	24	47	0	445
Total	828			

There will be many important uses for a good socio-economic scale. Some have felt that much of what is called intelligence is due to the nature of the home environment. It is interesting to note that in the investigation in which this scale was employed the raw correlation between mental age and educational age was reduced only from .739 to .701, by rendering socio-economic status constant.

I recommend that the scale be used in the following form:

QUESTION BLANK FOR PARENTS

Dear Parent:

We are making a study in which it is necessary to have the information asked for by this blank. The information will be used only in connection with our study. It will not affect the school standing of your child and it will not be seen by the teacher.

Will you answer the questions accurately and completely? Your coöperation will be much appreciated.

Yours truly,

.....

Name of child..... Date..... School.....

Item	Where answers are given draw a ring around right one	Do nothing with this column
1. Do you have a telephone in your home?	Yes No	17
2. Is your house heated by a furnace?	Yes No	16
3. How many books are there in your home?	.....	14
4. Write the name of your father's occupation	.....	13
(The name of your mother's if you have no father)	.....	13
5. Did the mother of the child attend high school?	Yes No	11
6. Did the father of the child attend high school?	Yes No	10
7. How many magazines do you take regularly in your home?	0 1 2 3 4 More	10
8. Not counting the bathroom, in how many rooms do you live?	1 2 3 4 5 6 7 More	8
9. Do you have a piano in your home?	Yes No	7
10. Do you have an automobile other than a truck?	Yes No	6
11. Do you have a phonograph or radio in your home?	Yes No	5
12. Is English the only language spoken in your home?	Yes No	4
13. How many daily papers do your take regularly in your home?	0 1 2 3 More	3
14. Do you live in an apartment house?	Yes No	

.....
(Name of Parent)

Note: Please put this question blank into the attached envelope. Seal the envelope and ask your child to give it to his teacher tomorrow morning.

Any occupation with a value of 11.17 and above on the Barr Occupation Scale should be credited with possession. Similarly, any responses indicating two or more newspapers, three or more magazines, six or more rooms, and 125 or more books should be credited with the possession value.

In the answer column, values higher than those used as the distinguishing points in the construction of this scale are given, because the blank might be used to collect data for a reconstruction of the scale. The score values of possession are given in the last column for convenience. It would probably interfere with the accuracy of the answers if the parent got an inkling of the meaning of these values. These values are crossed out for the omitted items and for those to which a response indicating non-possession has been made. The sum of the remaining values divided by the total number of items answered will be the score. The unit of the values given in the last column is ten times sigma.

SCHOLASTIC ACCOMPLISHMENT IN THE JUNIOR HIGH SCHOOL

F. C. LANDSITTEL
Ohio State University

THE present wide-spread agitation looking toward freeing the junior high school from the force of college entrance requirements is attended by at least one beneficent result. It is serving to center inquiry upon the effectiveness of this newer type of school as regards the laying of a proper foundation for scholarship. That weakness in this regard has been freely charged in the past, and that the charge is perennially renewed need not excite either wonder or protest. Being an experimental institution that runs counter to established conceptions, a certain amount of doubt must inevitably arise with regard to it. In other words, it is but the natural order of things that the burden of proof of the superiority of this innovation must be sustained by the innovation itself. Skepticism is not necessarily to be interpreted as evidence of hostility. Honestly entertained it may on the contrary serve to save the institution from its more radical friends.

The colleges, since they are asked to relinquish the part that they now exercise in determining the standards of the ninth grade, have a perfect right to require evidence showing that such a course of action will mean no material sacrifice of their own interests. And, since scholastic accomplishment represents their special interest, this is the point upon which the required evidence must center. Neither is the fact to be overlooked, in this connection that this factor of jealous concern to the colleges marks a prime excellence in all education, irrespective of whether it is designed specifically as preparation for college or not. Rigorous intellectual cultivation must be recognized as eternally vital if the general educational ideal ordinarily named by such terms as citizenship, culture, and character is to be realized in a reliable way.

In the interest, then, of the best development of the junior high school, there is presented in this article a comparison of the results achieved in three school branches by a total of 371 pairs of university freshmen. The course of preparation for college taken by one of the two individuals forming each of these pairs led through the three-year junior high school and three years in a senior high school; while that of the other led through the regulation eight grades and a four-year high school. Comparison of the records made by these students is limited to three subjects:

science, mathematics, and foreign languages. These subjects were chosen on account of the fact that in these fields are represented the widest departures of the junior high from the established order of teaching. It was thought, too, that in the case of the last some light might be thrown upon the question of whether or not there is advantage in beginning foreign language study below the ninth grade.

Preliminary to the process of pairing, a questionnaire was distributed in a limited number of freshmen groups through the courtesy of the English Department of the Ohio State University. In this way information was obtained regarding each of a total of 616 students, covering the following points: name of student; name and location of school in which preparation for college was completed; grades included, and location of junior high school attended, if any; years of attendance in junior high school; what science, if any, studied and name of textbook; what mathematics studied and name of textbook; and what foreign languages studied, if any. Such items as the name and location of the school attended, period of attendance, and subjects and textbooks studied were sought as helps in determining whether the junior high schools reported as attended were so in fact or only so-called. The investigator benefited very materially in this regard also from his familiarity with the schools of Ohio, attained through visitation of them covering a period of eight years as a half-time inspector of high schools.

Equivalent groups were obtained, as previously indicated, by the method of pairing. The primary consideration at this step was intelligence, the scores used being the percentiles in which the various students considered fell on the total distribution resulting from the entrance tests prescribed by the University for all freshmen entering in the fall of 1926. These scores were available through the courtesy of the Psychology Department of the University, which administers these tests. The factor of sex was disregarded. Ages were taken into account in pairing along with intelligence percentiles, the individuals paired being generally of the same ages and differences of more than one year being accepted in only rare instances. How close the comparison runs as respects these two factors of intelligence and age is shown by Tables I and II. The two individuals forming each pair we so selected, moreover, as to represent comparable social backgrounds to a larger extent than was believed to be possible when the investigation was started. In no case was a pair formed in which there was wide difference in either the sizes of schools or the types of communities represented.

For the science comparison 129 pairs were formed, all on the junior high school side having had general science. For mathematics 113 pairs were obtained, and for languages 66. In the case of mathematics, comparison of the full group could not be carried to the college level, owing to the fact that almost half on one side or the other took no mathematics in their freshman college year. Accordingly the comparison, so far as the full mathematics set was concerned, had to be limited to the high school level. After eliminating students not taking mathematics in college, the resulting 63 pairs were compared on both levels. The language groups obtainable turned out to be small also, owing to limited pursuit of the languages in college, and in high school as well.

The school records in the above named subjects for all students selected for comparison were next obtained. The necessary high-school grades were taken from the transcripts presented for entrance, which were made available through the courtesy of the University Examiner; and the college grades were secured from the individual student records on file in the offices of the Secretaries of the various colleges and the University Registrar, and kindly opened to our use by these officials. For obvious reasons the high-school records used were limited to the last three high-school grades, while on the college level the results of only the autumn and winter quarters of the academic year, 1926-1927, were available at the time the study was made.

In the large majority of cases a plurality of grades appeared in a particular field, as mathematics, which obviously had to be reduced to a single score. This was done by a process of averaging. On the high school level, for example, all tenth-grade marks in mathematics appearing in the high-school transcript were averaged, then all of the eleventh and of the twelfth, and the average of these averages was finally taken as the comparison score. College grades were similarly treated, the final marks recorded for all mathematics courses taken (if more than one) in the autumn quarter being averaged, and then this average being combined with a similar one for the winter quarter for the final average or comparison score. The complication presented by grades recorded in terms of letters was met by converting these into their numerical equivalents. This was facilitated by the fact that entrance transcripts are required to show the percent equivalents, and the further fact that a uniform rule in this regard is observed throughout the University. The middle of the percent span stated as the equivalent of any given letter was taken as the numeral entering into the averaging. For example, if a grade of B represented the span 85 to 94, then 90 was taken as this numeral.

The arithmetical mean of all individual averages in a general subject was computed for the final measure of each of the compared groups, the difference between the resulting two final measures being taken as representing the extent of superiority in point of scholastic accomplishment in the subject in question of one group over the other. Standard deviations and critical ratios were computed also as mathematical checks of reliability of the results obtained. The same statistical treatment was applied to intelligence scores and ages for the purpose of proving the reliability of the pairing. The importance to readers of this journal of the various distribution tables that were made is not regarded as sufficient to justify use here of the extent of space that would be required for their full presentation. Accordingly the complete tables are given only for science. These appear below, followed by two general summary tables.

TABLE I

INTELLIGENCE SCORES FOR COMPARED GROUPS IN SCIENCE

O. S. U. PERCENTILES	JUNIOR-SENIOR H. S. GROUP*	EIGHT-FOUR GROUP*
1– 5	8	8
6– 10	3	4
11– 15	6	6
16– 20	6	4
21– 25	5	6
26– 30	9	9
31– 35	8	7
36– 40	6	7
41– 45	6	7
46– 50	7	6
51– 55	10	10
56– 60	6	6
61– 65	10	9
66– 70	6	8
71– 75	7	6
76– 80	8	8
81– 85	6	6
86– 90	2	2
91– 95	7	7
96–100	3	3
Total	129	129
Mean	49.5	49.0
SD	25.5	26.5
SDM	.198	.205

TABLE II

THE CHRONOLOGICAL AGES OF GROUPS COMPARED IN SCIENCE

AGES	JUNIOR-SENIOR H. S. GROUP	EIGHT-FOUR GROUP
15	1	
16	6	2
17	29	24
18	48	43
19	28	32
20	11	14
21	6	9
22		2
23		0
24		1
25		1
26		0
27		1
Total	129	129
Mean	18.2	18.7
SD	1.2	1.5
SDM	.009	.012

*In this and following tables the designation *Junior-Senior High School Group* is used to refer to a paired group whose pre-college training embraces three years of attendance in a junior high school of the three-year type and three years in a senior high school; while the *Eight-Four Group* refers to a paired group composed of pupils with the conventional eight-grade and four-year high-school course of preparation.

TABLE III

DISTRIBUTION OF GRADES MADE IN THE SENIOR HIGH SCHOOL BY MEMBERS OF THE JUNIOR-SENIOR HIGH SCHOOL AND OF THE EIGHT-FOUR GROUP

GRADES	JUNIOR-SENIOR H. S. GROUP	EIGHT-FOUR GROUP
63–67	2	1
68–72	8	5
73–77	23	24
78–82	24	29
83–87	31	23
88–92	23	23
93–97	17	24
98–	1	
Total	129	129
Mean	83.4	84.1
SD	7.8	7.6
SDM	.68	.67
D=.7	SDD=.95	E. C.=.77

TABLE IV

DISTRIBUTION OF GRADES MADE IN THE FRESHMAN YEAR OF COLLEGE BY MEMBERS OF THE JUNIOR-SENIOR HIGH SCHOOL AND OF THE EIGHT-FOUR GROUP

GRADES	JUNIOR-SENIOR H. S. GROUP	EIGHT-FOUR GROUP
58–62	10	1
63–67	5	6
68–72	14	16
73–77	22	19
78–82	35	39
83–87	13	15
88–92	16	13
93–97	14	20
Total	129	129
Mean	79.3	81.1
SD	9.7	8.7
SDM	.85	.77
D=1.8	SDD=1.15	E. C.=.56

TABLE V

INTELLIGENCE RATINGS AND AGES OF PAIRED GROUPS

BASES	GROUPS	N	M	SD	SDM
	SCIENCE				
Intelligence Ratings	Jr.-Sr. H. S.	129	49.5	25.5	.198
	Eight-Four	129	49.0	26.5	.205
Ages	Jr.-Sr. H. S.	129	18.2	1.21	.009
	Eight-Four	129	18.7	1.5	.012
	MATHEMATICS (Large Group)				
Intelligence Ratings	Jr.-Sr. H. S.	113	52.5	20.0	.18
	Eight-Four	113	50.5	27.0	.24
Ages	Jr.-Sr. H. S.	113	18.0	1.34	.012
	Eight-Four	113	18.5	1.18	.01
	MATHEMATICS (Small Group)				
Intelligence Ratings	Jr.-Sr. H. S.	63	59.5	26.0	.41
	Eight-Four	63	59.0	26.0	.41
Ages	Jr.-Sr. H. S.	63	17.9	1.29	.0204
	Eight-Four	63	18.4	1.4	.022
	LANGUAGES				
Intelligence Ratings	Jr.-Sr. H. S.	66	52.0	19.0	.28
	Eight-Four	66	52.0	19.0	.28
Ages	Jr.-Sr. H. S.	66	18.1	1.3	.02
	Eight-Four	66	18.3	1.1	.017

NOTE.—The above tables are read as follows: The intelligence ratings of students with junior-senior high school preparation used in the science comparison numbered 129; they have a mean of 49.5; the standard deviation is 25.5; and the standard deviation of the mean is .198

TABLE VI

A COMPARISON OF THE MEAN PERFORMANCE (GRADES) OF THE JUNIOR-SENIOR HIGH SCHOOL AND THE EIGHT-FOUR GROUP

LEVELS	GROUPS	N	M	S.D.	S.D.M.		
SCIENCE							
Sr. H. S.	Jr.-Sr. H. S.	129	83.4	7.8	.68	D = .7	
	Eight-Four	129	84.1	7.6	.67	SDD = .95	Critical ratio = .74
College	Jr.-Sr. H. S.	129	79.3	9.7	.85	D = 1.8	
	Eight-Four	129	81.1	8.7	.77	SDD = 1.15	Critical ratio = 1.57
MATHEMATICS							
Sr. H. S.	Jr.-Sr. H. S.	113	83.6	7.3	.68	D = .8	
	Eight-Four	113	84.4	6.6	.62	SDD = .92	Critical ratio = .87
Sr. H. S.	Jr.-Sr. H. S.	63	84.7	6.6	.83	D = 1.1	
	Eight-Four	63	85.8	6.4	.8	SDD = 1.15	Critical ratio = .96
College	Jr.-Sr. H. S.	63	77.6	8.9	1.12	D = .9	
	Eight-Four	63	76.7	10.4	1.31	SDD = 1.4	Critical ratio = .64
LANGUAGES							
Sr. H. S.	Jr.-Sr. H. S.	66	81.4	6.7	.81	D = 1	
	Eight-Four	66	82.4	8.8	1.08	SDD = 1.35	Critical ratio = .74

NOTE.—Read as follows: The general average in science of the 129 students having junior-senior high school preparation is 83.4; the standard deviation, 7.8; and the standard deviation of the mean, .68. For the 129 students having the eight-four type of preparation the mean is 84.1, the standard deviation, 7.6, and the standard deviation of the mean .67. The difference of the two means is .7 in favor of the latter group; the standard deviation of the difference is .95; and the critical ratio, .74. being $\frac{D}{SDD}$

Before attempting to formulate the conclusions to which the statistical evidence presented in the foregoing tables points, it will be well to take account of certain limitations attaching to the original data employed in this study. Questionable factors are involved in the bases of pairing as well as in the measures of student accomplishment.

With reference to the first of these, the fact is to be observed that both the intelligence ratings and the ages are those obtained at the time of entrance into the University. Whereas parity of ages as between paired individuals obviously has been constant throughout the period covered by the study, this cannot be presumed with certainty in the case of intelligence. In other words, we cannot say positively that two individuals rated as of equal intelligence at the time of entering college were so three years earlier, and throughout the intervening period. Abundant evidence warrants a strong presumption in that direction, however. There is not as a rule any considerable fluctuation in the intelligence of individuals, excepting as they may be affected by abnormal physical condition. Such fluctuations are, however, much less likely to occur in post-pubescent years, the period of development with which this study is concerned, than in pubescent or pre-pubescent years. The two main

criteria with reference to which our equivalent groups were formed appear, therefore, to be at least reasonably valid. This accepted, the relative exactness of the pairing becomes apparent when the means and deviations of paired groups are referred to.

Then there will be those who will not be disposed to accept the final semester grades in high school and college as reliable measures of pupil accomplishment. While it must be admitted that the subjective judgments of teachers enter into these final ratings of accomplishment, their reliability is much less open to question than ordinary or occasional marks put upon students' work. They represent mature consideration, are based upon considerable accumulations of evidence in the usual case at least, and have in some considerable measure been influenced by objective forms of testing. A probably more serious limitation attaching especially to the high-school grades that have been used in this study is the variation in standards of marking, presumably obtaining among the numerous schools whose graduates are represented in the study. Just how vital the effect of this variation may be, there appears to be no means of knowing. When taken together with the limitations previously stated, it probably means that our compared measures are at best only approximations. The claim may legitimately be made, however, on the other hand, that these final grades form at least the best available designations for extent of general accomplishment in the several subject-matter fields.

The limitations just pointed out as attaching to our original data must be regarded as weighty enough to make the results inconclusive, whatever may be the attitude that one entertains toward the mathematical method applied. The defects involved seem, however, not to be serious enough to render the outcome valueless. With these facts in mind the results may be briefly summarized as follows:

1. In the field of science the scholastic accomplishment in the senior high school of 129 students with junior high school training preparation is represented by an average score of 83.4 on a scale of 100. The corresponding average for an equivalent ability group finishing their college preparation in the conventional four-year high schools, with preliminary training in an eight-year elementary school, is 84.1. The difference in these two averages is .7 in favor of the eight-four group. This difference is, however, not statistically significant as shown by the critical ratio, .74. These data mean that the chances are seventy-seven to one hundred, upon indefinite repetition of the comparison, that a difference above zero will appear in favor of the eight-four organization.

2. The average score in science of the same junior-senior high-school group for the first two quarters of their freshman year in college is 79.3, and that of the opposed group, 81.1. The difference in the average score for the two groups is 1.8, in favor of the eight-four group, the chances being ninety-four out of every hundred of recurrence with repetition of the comparison of a positive difference favoring the same side.

3. In mathematics the large junior high-school group, numbering 113, made on the high-school level an average score of 83.6, and the eight-four group 84.4. The difference in this case is .8 in favor of the eight-four group, as formerly, the chances of recurrence of a positive difference favoring that side being eighty-one out of every hundred.

4. The small junior high-school group, numbering 63, averaged in their later work in high-school mathematics 84.7 and the eight-four group 85.8. The difference here is 1.1, again in favor of the eight-four group, and is significant enough to show chances of eighty-three out of every hundred.

5. The respective averages of these same two groups in college mathematics are 77.6 and 76.7. The difference in this case being .9, thus shifting to favor the junior-senior type of organization. The chances indicated in this case are seventy-three out of every hundred.

6. In languages, with regard to which comparison was found not to be feasible beyond high school, the averages are 81.4 for the junior-senior high-school group for which there were 66 cases and 82.4 for the eight-four group. The difference is 1, and in favor of the eight-four group, with chances of seventy-seven out of every hundred. The number of cases here is so small as to militate seriously against the reliability of the result.

7. The general showing is not to the credit of the junior high school; every really significant difference is against it. The comparison of the small mathematics group is the only one which does not agree with the general trend of these results. This discrepancy is, however, very largely minimized by the fact that it is based upon only sixty-three cases, and by the further fact that its reliability, as indicated by the critical ratio in this instance, was the lowest of all.

It is pertinent to refer in this connection to another very recent investigation reflecting upon the effectiveness of the junior high school as regards the fostering of scholarship. It is a survey of the junior high schools of Minneapolis, Minnesota, reported within the current year by J. Orin Powers in a University of Minnesota monograph. Standardized

tests in reading, arithmetic, geography, and American history were given in grades VII-B to VIII-A of junior high schools and of schools of the more conventional type, with results of doubtful credit, if at all so, to the newer form of organization. While the surface showing from these tests is distinctly unfavorable; when other factors, principally native intelligence are considered the differences are not marked. The general outcome of the survey can perhaps be set forth most satisfactorily by quoting from Dr. Power's monograph. He summarizes on this point as follows:

> "There is little evidence in the study that through reorganization better individual scholarship has been attained. There is evidence, however, that superior advantages may be gained by reorganization through better retention, more rapid promotion, enrichment of curricula, etc., without a loss in individual scholarship. The evidence both from random sampling and from group comparisons is indicative of slightly better individual accomplishment in reorganized schools."

Other studies have resulted in nothing more conclusive than the study by Powers. Taken together these investigations are sufficiently in agreement to make it impossible to maintain successfully the superiority of the junior high school as it affects individual scholarship. In other words, the fact seems to be that the new school has failed thus far to justify itself in regard to the central objective of education; viz., intellectual cultivation. This is deeply disappointing. There is reassurance to be found, fortunately, in the broader and richer culture that it affords; accordingly one's faith in reorganization may be preserved while the weaknesses brought out by this investigation are corrected.

MENTAL ABILITY WITH REFERENCE TO SELECTION AND RETENTION OF COLLEGE STUDENTS[1]

F. P. OBRIEN
University of Kansas

MANY and conflicting claims have appeared in educational literature concerning the importance of mental ability as a factor in determining student success in college. Sentiment has frequently been expressed to the effect that the towering increases in college enrollments of recent years are accompanied by grave dangers to both the students and the institutions. On the one hand, it is claimed that many more students are entering college than have the ability to profit by college opportunities, thereby resulting in disappointment, failure, and economic loss to them. On the other hand, the claim is made that the normal functions of the institution are impeded by congestion in enrollments and that, while the cost of its operation is greatly increased, at the same time the quality of its output is being reduced. As a consequence it seems that college administrators have become increasingly concerned with keeping out the unfit—a negative type of selection. The present study is concerned with an inquiry pertaining to how successfully colleges attract and hold 'the fit' and how successfully they appear to minister to those attracted.

As a first step in this study, the Terman Group Test of Mental Ability was administered to more than four thousand students enrolled in the junior or senior year of high school. Approximately sixty high schools (in Kansas) were represented, whose total (4-year) enrollments ranged from 25 to 750 students. The tests were given under definite direction and the scoring was done under close supervision. That the scores were fairly reliable was indicated by their correlation with the scores secured later by repeating the test in several of the schools. For the students who were 14 or 15 years of age at the first testing the Pearson r between the tests was $.96 \pm .13$; for those 16 or 17 years of age the r was $.90 \pm .025$. These two age-groups comprised 80 percent of the age distribution involved. Further evidence of their reliability was provided by correlating the Terman scores in some schools with scores in the Otis Self-Administering Test, and by correlating the scores of those students who later entered the University of Kansas with scores made in the mental test given to entering students.

[1] This paper in abbreviated form was read before Section Q of the American Association for Advancement of Science, at Nashville, December, 1927.

As a second step, by securing the close and continued coöperation of school officials in the schools in which these high-school students were enrolled, it was possible to follow them through high-school graduation and also to trace those who entered college to the collegiate institutions in which they enrolled. Similar facts were secured concerning even those whose high-school graduation was delayed or who delayed for a year or two after graduating before entering college. Next, the college record of these students was made the subject of careful inquiry. This study has now extended over a period of more than six years.

TABLE I

HIGH-SCHOOL GRADUATES AND COLLEGE SELECTION WITH REFERENCE TO MENTAL TEST SCORES

Mental Scores at H. S. Testing	Number of Seniors Tested in High School	Number Seniors Graduated	Graduates Who Went To College	Graduates Who Did Not Go To College	Number of Juniors Tested In High School	Number of Juniors Graduated	Graduates Who Went To College	Graduates Who Did Not Go To College
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
211–220	2	2	2					
201–210	20	18	10	8	7	6	3	3
191–200	38	37	19	18	31	27	17	10
181–190	109	104	54	50	62	53	24	29
171–180	125	119	51	68	123	111	63	48
161–170	164	157	59	98	134	112	51	61
151–160	190	182	71	111	167	136	73	63
141–150	189	171	71	100	193	153	70	83
131–140	166	159	59	100	265	225	74	151
121–130	189	174	55	119	246	203	63	140
111–120	163	152	51	101	220	180	47	133
101–110	137	127	30	97	197	160	47	113
91–100	77	69	19	50	164	126	22	104
81– 90	54	49	5	44	109	82	19	63
71– 80	37	31	5	26	88	60	11	49
61– 70	16	10	2	8	50	36	5	31
51– 60	10	8	1	7	22	16	2	14
Below 51	2	1	1		14	8	1	7
Totals	1688	1570	565	1005	2092	1694	592	1102
Medians	140	141	149	136	129	130	142	124
25 Percentile	116	117	126	113	105	107	120	102
75 Percentile	163	164	170	160	151	153	163	144

Thirty-five percent of the students who graduated from high school entered college. This number was exclusive of students who entered so-called colleges of business, nursing, osteopathy, music, et cetera. Composite tabulations were made of the mental test scores of the students tested in the various high schools. Analysis of these tabulations disclosed (1) that the median scores were a little below the Terman norms (8 points), (2) that 40 percent of the students who did not finish the high-school course ranked above their class average, (3) that 61 percent

of those who entered college had mental scores which placed them in the upper fifty percent of their high-school graduating class. The assumption that the upper half of this composite group of high-school graduates represented a definite degree of selection with reference to mental ability may be generally accepted.

Due to the fact that some high schools which coöperated in giving the tests did not continue their coöperation in the subsequent portions of the study, the tabulations include only the 3,780 students for whom the information was complete. Of this number 1,688 were tested as seniors and 2,092 as juniors.

It may be considered a fact of significance that only 44 percent of those high-school graduates whose scores in the mental ability test were above the average have presented themselves for admittance to any college. Consequently there were more of these relatively superior students (56 percent) who did not enter any college than the number of that type who did enter college. It is appropriate to observe here that the institutions in the state which seemed by their locations to provide easily available opportunities for these students included a state university, three teachers colleges, a state agricultural college, various denominational colleges and a dozen public or private junior colleges.

Whatever may be the explanation, it is evident that a majority of those students who are presumably best fitted by mental ability to receive training in higher institutions were not candidates for such training; and that 40 percent of those who did enter college had a mental rating below the average of the composite group of high-school graduates of which they formed a part. Certainly the question invites more than passing attention as to why college opportunities fail to attract a majority of the students of better ability from these high schools? This study presents the situation but offers no solution to that question. However, it is evident that more than twice as many of these students might have entered college without reducing the mental caliber already represented by those who did go to college. It also becomes evident in Table II that only half of those who were above average in mental ability among students entering college have graduated from college. (Some of the 21 now in college may also graduate.)

To make sure that the high-school graduates who were reported as entering college actually matriculated in college the records of the registrar's office and the college directories were investigated. Later a detailed study of college records provided full information concerning the achieve-

ment and progress of these students in several colleges. The study was restricted for practical reasons to those attending nine institutions, all located within the state. Of this number three were state institutions—the State University, the State Agricultural College, and the Emporia State Teachers College, and six were privately controlled colleges. The latter included Baker, Ottawa, Washburn, McPherson, Emporia, and Bethany. This analysis of college records involved 722 students who seemed from a comparison of mental test scores to represent a fair sampling of the total number who matriculated in college.

TABLE II

MENTAL TEST SCORES WITH REFERENCE TO RETENTION OF STUDENTS IN COLLEGE

MENTAL SCORES AT H. S. TESTING (as seniors)	NUMBER WHO ENTERED COLLEGE	COMPLETING THE AMOUNTS OF COLLEGE WORK INDICATED				GRADUATING FROM COLLEGE*
		Less than one year of college	One year but less than 2 years	Two years but less than 3 years	Three years but not completing course	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
211–220	4		1	1	1	1
201–210	17		5	2	1	9
191–200	28		2	4	2	20
181–190	74	2	12	16	6	38
171–180	70	1	9	14	5	39
161–170	85	8	14	16	4	40
151–160	72	6	16	13	5	30
141–150	95	8	21	25	3	34
131–140	87	6	23	17	5	33
121–130	56	5	17	18	3	11
111–120	57	8	16	11	4	16
101–110	36	4	10	6	1	14
91–100	18	4	6	3	0	5
81– 90	11	1	3	2	2	3
71– 80	8	2	2	2		0
Less than 70	4		2	1		1
Totals	722	55	159	151	42	294
Medians	150	137	137	147	157	161
Middle 50%	129–173	115–156	120–163	129–170	132–180	138–179

*Eighty students still in college who are reasonably assured of graduating have been classed with those graduating. Twenty-one others are still pursuing the first half of their college work.

Table II records 55 withdrawals from college before the end of the first year and 159 others before the end of the second year of college work. A total of 407 have withdrawn and others of the number who are in college may withdraw later.

The mental scores of the two classes tested in high school have been combined into a single table here for the sake of simplification. For this purpose the scores of junior students were 'stepped-up' by an amount equal to the average difference between the scores of college entrants from the two classes. The scores were then regarded as though made by high-school seniors and combined with those actually made by seniors.

More students drop out of college than graduate from college. The eliminations according to official college records constituted 57 percent of these 722 entrants. However, this is an understatement since not all of the others have as yet completed their work for a degree. For convenience in reporting the facts, the graduation of 80 students who are now well advanced in college has been anticipated and they are reported with the graduates. A few of them may fail to graduate, and 21 other students who are still in college have not yet completed half of the college course. Forty-two percent of those eliminated from college ranked in mental test scores above the average of students who entered college. On the other hand, 40 percent of those who continued in college to graduation did not score above the average of college entrants in mental ability.

Of the students who withdrew from college only 47 percent continued their work as long as two years. Need for complete and careful diagnosis of the situation pertaining to the first years of college seems evident. Many students apparently competent in ability and presumably serious of purpose in starting college are soon lost from college and its announced opportunities. The charge that these students lack the required degree of mental ability for success in college work seems to be insufficient as an explanation. It may be pointed out in this connection that from the upper 50 percent of high-school graduates less than one in four graduated from college.

Some students withdraw from college for the purpose of transferring to another institution and others who withdraw return to the same college. If the student returned to college after withdrawing he was not reported in this study as withdrawing. Twenty-two students who later enrolled in other colleges in the state were not reported in this study as withdrawing from college if the new enrollment continued to graduation. A definite effort was made to trace all students who shifted from one college to another. Some were doubtless overlooked. Especially would this be true if any students transferred to colleges outside of the state. Whether such transfers occurred or how frequently no information seemed to be available. It is assumed that they were unusual.

Records of the scholastic grades made by these students in college were also given consideration. The four passing grades (A, B, C, and D) were assigned respective values of 4, 3, 2, and 1. All grades of each individual were translated in accord with this plan and the total of these values divided by the total of hours credited to him in order to secure an

average credit-hour grade in college for the individual. The mean scholastic grades thus secured are shown in Table III.

TABLE III

MEAN SCHOLASTIC GRADES EARNED IN COLLEGE

	MEAN GRADES								TOTALS	MEDIAN GRADE
	0	1	1.5	2	2.5	3	3.5	4		
Graduating students			20	80	94	63	25	12	294	2.5
Students who withdrew from college	23	23	101	134	89	26	10	1	407	2.0

The average grade made by 12 graduating students was 4 (equivalent to an A). Twenty-five others made a mean grade about midway between an A and a B (3.5). The students who withdrew may be represented scholastically by a C (2), and those graduating, by a half-way grade between C and B (2.5). The 23 students who are designated as having 0 grade took no tests to provide the basis of a grade. The extent to which lack of credit was due to failing grades is shown in Table IV.

TABLE IV

THE NUMBER OF HOURS OF 'F' GRADES RECEIVED IN COLLEGE

	HOURS OF F's								
	0	1–3	4–6	7–9	10–12	13–15	MORE THAN 15	TOTALS	MEDIAN NO. OF HOURS
Graduating students	179	31	19	15	19	13	18	294	8.5
Students who withdrew from college	240	29	41	25	21	13	38	407	9.3

Almost exactly the same percentage of the students who withdrew from college and of the students who continued in college (40 percent) had F's recorded in their college grades. Furthermore, the average number of hours of F's recorded was markedly similar for members of these two college groups. The 115 graduating students who had F grades had an average of 8.5 hours of F's. Bearing in mind that 80 of these graduates are near graduation rather than actually graduated, there is a possibility of some additional F's to increase the average as stated here. The 167 students among those withdrawn who had F grades failed in an average of 9.3 hours of work. Of course, the latter group were in college approximately half as long a time as were those who graduated. One does not know to what extent the F grades influenced withdrawing from college or to what extent both the F's and withdrawing resulted from

other factors which operated as causes. Yet it is evident that 18 students who failed in the equivalent of at least one semester of college work (more than 15 hours) still were able to meet the requirements for graduation.

Assuming the existence of some definite relation between the quality of scholastic work in college and the time devoted to non-scholastic interests or to self-support activities, an earnest effort was made to secure the two kinds of information indicated from the several institutions whose records were consulted. Perhaps one's success in college may be affected even more by such influences than by his ability as measured by a mental test. The effort was not very successful due to the incompleteness or absence of the data needed. However, the college records of 160 of the high-school graduates involved in this study were traced through two years of college work and in four different institutions. This work was done by Miss Lou LaBrant in a Masters thesis study.[2] One of the conclusions reached in that study is quoted here:

"Although college students who ranked in the lowest quartile as measured by the Terman mental test (given in high school) failed in three times as many hours of credit as those in the highest quartile, they carried extra-curricular activities (on a point basis) in the ratio of 2 to 3 in comparison with students in the highest quartile and were self-supporting (on a point basis) in the ratio of approximately 4 to 1 as compared with the highest quartile students."

The facts presented in this paper may seem at variance with claims which are frequently made and with the conclusions of some other studies. Yet there is evidence abroad of a gradually growing conviction that inducing students to use effectively the ability which they possess may frequently be more practical and fruitful than ever looking for better material. It was not the province of this study to inquire what influence was exerted on the selection and retention of these college students by the content of college instruction, by the prevailing method of measuring achievement in college, or by the effectiveness of high-school instruction as a preparation for college although the need of such inquiries as these is at least suggested.

Summarizing this study, more than 4,000 students who were enrolled in the upper classes in 60 high schools were tested; from most of the schools information was secured concerning the high-school graduation of

[2] LaBrant, Lou. "Intelligence Rating of High School Pupils and Their Achievement in College." University of Kansas, 1925.

these students and their subsequent enrollment in college; and detailed information was obtained from the records of nine colleges, over a period of several years, concerning the graduation or elimination as well as the scholastic success of the students who matriculated in those colleges. The study has disclosed that 35 percent of the high-school graduates entered college; that 61 percent of those who entered college came from the upper 50 percent of high-school graduates; that fewer than half of the high-school graduates who ranked above average in mental ability went to college; that fewer than 1 in 4 of this portion of high-school graduates became college graduates; that the majority of the students who entered college withdrew before graduation—more than half of the withdrawals came before the end of two years in college; that 42 percent of the students eliminated from college ranked in mental ability above the average of students who entered college; that the mean scholastic grade made by the students who remained in college until they graduated was midway between a C and a B, while the mean grade of those who withdrew was a C; that the same percentage of students made F's in both groups; and that the average number of hours of F's by failing students did not differ greatly for the two groups.

There is some evidence that students in college who rank low in a mental test are self-supporting to a much greater extent but participate in non-curricular activities to a lesser extent than do students who rank high in such a test. Finally, it appears from the facts at hand that the colleges are really successful neither in attracting nor in holding the mentally fit.

WASTE IN PROFESSIONAL EDUCATION

RICHARD E. HYDE
Fairmont State Normal School

EVERYBODY who has considered the matter at all realizes that many, and very probably all, social sciences have a large amount of overlapping and duplication of subject matter. Repetition of subject matter causes loss of interest on the part of the student, and a feeling that one's time is being wasted. One hears the accusation of "snap" course against the social sciences more often than against the natural sciences. Most certainly a course in college may be easy, and still be very much worth while. But if the course is easy, because a student already knows a considerable part of it, this is a situation which merits attention, and remedial measures should be taken at once.

Education is one of the college subjects in which the amount of overlapping and duplication of subject matter has received some attention. Certain solutions to the problem have been suggested. Anent this situation, in the February, 1926, issue of *Educational Administration and Supervision*, in his article entitled "Twice-Told Tales, or Tautologous Teaching in Higher Education,"[1] Dean A. Worcester, of Ohio University, summarizes his findings as follows:

"In two institutions students just beginning the introductory course in Educational Psychology were given examinations designed as final examinations for the course. In each case it was found: (1) That nearly all of the students knew something of the subject-matter of the course; (2) that a large proportion of the students gave evidence of knowing a good deal—from 20 per cent to 40 per cent—of the course; (3) that certain students had in some way or other acquired so much of the matter of the course that it was very doubtful if they should be allowed to go on with it; (4) that certain topics were already familiar to everybody."

The following investigations, carried on at the Fairmont State Normal School, Fairmont, W. Va., are similar to the aforementioned, but were made with the purpose of finding out how much worse, if any, the situation was with upper classmen than with freshmen. It is reasonable to suppose that the repetition and overlapping of subject matter is worse in the last two years of college than it is in the first.

[1] Also published as one of the chapters of *Research Adventures in University Teaching*, Bloomington, Illinois: Public School Publishing Company.

FIRST INVESTIGATION

In February, 1927, at the start of the second semester, thirty-four junior normal students, or college freshmen, beginning a course in elementary educational psychology were given a test to determine their present knowledge of the field. The test, consisting of three items of the true-false variety, was intended to be a final examination in the subject. The papers were scored by subtracting the number wrong from the number right. The median score was 23.1+. The range of scores was from −5 to 54. The results of the test were tabulated item by item, and it was found that the percentages of correct answers to each of the items were as follows: 77, 66, 80, 57, 74, 31, 79, 60, 54, 91, 66, 90, 75, 90, 83, 50, 97, 40, 71, 79, 71, 79, 46, 26, 71, 71, 46, 12, 54, 63, 80, 83, 71, 74, 80, 54, 9, 51, 47, 83, 41, 74, 96, 80, 97, 51, 3, 46, 9, 60, 63, 54, 57, 66, 74, 30, 57, 90, 90, 60, 83, 60, 57, 34, 33, 33, 54, 40, 71, 51, 74, 43, 46, 75, 83, 34, 79, 29, 83, 71, 50, 57, 74, 66, 30, 51, 46, 63, 74, 66, 63, 46, 83, 66, 20, 43, 83, 40, 51, 40, 74, 60, 33, 51, 43, 80, 14, 37, 54, 67, 23. By scanning these percentages it is seen that 28 of these items were known to 75 percent, or more, of the students; seventy-eight of these items were known to 50 percent, or more, of the class. All but three members of the class got at least half of the statements correct. The three members who did not get one-half of the items correct came within a few points of doing so, the lowest number right being 53. In terms of the test results, each student was familiar with approximately one-half, or more, of the course before it started.

At the end of the semester a similar test with the same number of items was given. The median score was 47.5. If we adopt for our standard of marking the curve of chance, assigning failing marks to one-tenth of the class, the lowest passing score was 37. In the first test alluded to above, six students made a score as good as, or better than 37. In other words, at the beginning of the course, six out of 34 students already had a knowledge of the subject which would ordinarily have given them credit for the course. Five of these had scores on the first test of 35-39, while one made a score of 51. It is true that all except one of this group made much better scores on the final test, so seemingly their time was not entirely wasted, but undoubtedly a large part of the course was an old story to them.

SECOND INVESTIGATION

The second investigation was carried on during the summer session of 1927 at the Fairmont State Normal School with junior and senior college students. A class of fifteen, enrolled in a course in high-school methods, was given a true-false test of 100 items. This preliminary testing showed a median score of 31.7—. The range was from 7 to 53. The test was scored item by item. A summary of the percentage of correct answers is given in Table I.

TABLE I

PERCENTAGES OF CORRECT ANSWERS

PERCENTAGES OF CORRECT ANSWERS	FREQUENCY
100	7
94	7
87	9
81	10
75	13
69	13
63	1
62	9
56	3
50	5
44	6
37	6
32	1
31	2
25	2
19	4
14	1
0	1
Total	100

Seven of the items were known to everybody, 7 were known to 94 percent of the group, 9 were known to 87 percent of the group, etc. The correct answers to 46 of the statements were known by three-fourths of the class at its inception, while one-half of the class was familiar with at least 77 items. In considering the number of statements correctly answered by each member of the class, it was observed that all answered correctly over half of the items, the lowest number right being 53. Two had 76 correct, indicating that about three-fourths of the course would be a "Twice-told Tale" to them. These people had had anywhere from 12 to 23 "college hours" of education in previous courses.

A final examination of the same type and similarity of informational content was given at the end of summer school. In this case the median was 51.7—. The range was 24 to 78. If the lower ten percent of the class is taken as failure, it is found that two had a passing knowledge of the course at its beginning.

COMPARISON OF THE RESULTS

Comparing the results of the two investigations, it is noticed that six out of 34 freshmen and two out of 16 upper classmen had a passing knowledge of the course at the time they enrolled for it. The percent in the former case is 17+, and in the latter, 13−, showing in this particular that conditions are slightly worse with the underclassmen. But a detailed analysis of the test results is more indicative of the true condition of affairs. Seven items were known by all the upper classmen; not a single item was familiar to all the underclassmen. Twenty-eight items were known by three-fourths of the freshmen, but three-fourths of the juniors and seniors knew 46 of their 100 statements. One-half of the upper classmen knew 77 items, or more, while the other group knew 76, the two being almost equal in this particular. The conditions are worse with the upper classmen, but it is surprisingly bad with the underclassmen when one considers that the latter are beginners.

One objection that may be made to the findings of these investigations is that not enough subjects took part in them. However, if these conditions are true of small sections, with large sections conditions are much aggravated, for here you have possibilities of wider ranges of individual differences in achievement and other particulars. Although the sections are small, there is no reason to believe that they are not representative of general tendencies. With a larger number of subjects (185 in this case) Dean A. Worcester found that 22 percent received a score equal to or greater than one which would be accepted as a passing mark at the conclusion of the course. (See article previously referred to.) In this case at least the conditions were worse than those depicted in the preceding paragraphs, but it is only a comparative matter. All these investigations show the paramount necessity of carefully thought out plans to counteract the conditions revealed.

Many different suggestions to rectify these findings have already been made. Some of these are (1) giving a test at the beginning of the course, and segregating certain groups, using the test results for this homogeneous grouping, or if the group is not large enough for this treatment, divide the class itself into several sections for such work as assignments, the purpose always being to meet in a better way the needs of those who already have a previous knowledge of the course; (2) examining the syllabi of the different courses, and having some agreement as to what course certain topics belong in. Some topics are found in at least three

different or supposedly different courses, with similarity of treatment. Perhaps some of these topics belong in all three, or two, or four, or whichever is the number of courses they are found in, but undoubtedly not as many belong there as are commonly found.

SUMMARY

In the Fairmont State Normal School a class in educational psychology composed of freshmen, and a class in high-school methods composed of juniors and seniors were given tests at the beginning of the course to find out their present knowledge of the course. It was found (1) that all of the members of each course displayed some knowledge of the course; (2) that the upper classmen as a whole were more familiar with the instructional content to follow than the freshmen were with their prospective subject matter; (3) certain segments of information were known by almost every one; (4) that each upper classman knew half or more of the true-false statements in his test; (5) that each of the freshmen possessed knowledge of almost one-half or more of the course.

RESEMBLANCE IN THE HANDWRITING OF TWINS AND SIBLINGS

EMILY KRAMER
University of Washington
AND
CHARLES E. LAUTERBACH
Harvard University

THIS study concerns itself, as the title of the article indicates, with similarities in the handwriting of twins and siblings. The writers wished to determine, first, whether twins show a greater degree of resemblance in handwriting than siblings; second, whether like-sex pairs of twins show a greater degree of resemblance in handwriting than unlike-sex twins; third, whether older twins show a greater degree of resemblance in handwriting than younger twins; and, fourth, whether the presence of symmetry reversal in twins affects the degree of resemblance in handwriting.

Data were collected on 205 pairs of twins in the public schools of Wisconsin and 101 pairs of siblings in the public schools of Montana. The ages of both groups ranged from seven to twenty years. To test these pupils, the Kansas City Scale for Measuring Handwriting, Missouri Edition, was used. The scoring of each paper was checked twice for rate and four times for quality by three different judges, who are designated as A, B, and C in Table I which gives the reliability coefficients between the different judgments of quality on the handwriting of the two groups. Judge A scored the papers twice.

TABLE I

COEFFICIENTS OF CORRELATION BETWEEN THE QUALITY SCORES OF THREE JUDGES ON THE HANDWRITING OF TWINS AND SIBLINGS

JUDGES	SIBLINGS	TWINS
(1)	(2)	(3)
A-A	.882 ± .010	.917 ± .005
A-B	.856 ± .013	.886 ± .007
A-C	.844 ± .014	.813 ± .011
B-C	.873 ± .011	.831 ± .010

The similarities in handwriting for the various groups as expressed by coefficients of correlation are shown in Table II. Here the symbols in Column 1 indicate the special groups or pairs; a-b, brother-brother pairs; x-y, sister-sister pairs; a-x, brother-sister pairs. The letters "SR" represent "symmetry reversal."

TABLE II

COEFFICIENTS OF CORRELATION BETWEEN THE SCORES OF VARIOUS GROUPS OF SIBLINGS AND TWINS AS MEASURED BY THE KANSAS CITY HANDWRITING SCALE

GROUP	NUMBER OF PAIRS	QUALITY	RATE
(1)	(2)	(3)	(4)
	Siblings		
a-b	25	.46	.27
x-y	32	.24	.09
a-x	44	.13	.16
	Twins		
a-b	69	.69	.82
x-y	74	.68	.84
a-x	61	.37	.37
abxySR	32	.65	.82
axSR	29	.41	.31
Younger	111	.49	.66
Older	93	.58	.55

Pearson, in 1904, found correlation coefficients for the handwriting of siblings ranging from .48 to .56.[1] His correlation for the a-b Group was .53; for the x-y Group, .56; for the a-x Group, .48. His scores were based entirely on teachers' judgments and not on scientifically constructed scales. Starch has reported the results of two investigations of handwriting made some ten years ago.[2] In 1915, he found correlation coefficients of .06 for quality and .18 for rate; in 1917, .46 for quality and .72 for rate. Galton, in his book on *Inquiries into Human Faculties*, discussed the handwriting of twins. He stated that the one item in which similarity was rare was handwriting.[3]

The correlations presented in this study, however, indicate considerable resemblance in the handwriting of twins. Lauterbach recently offered an explanation for this seeming conflict with Galton's opinion.[4] He

[1] Pearson, Karl. "On the Inheritance of Mental and Moral Characters in Man and Its Comparison with the Inheritance of Physical Characters." *Biometrika*, III (March and July, 1904), 131-90.

[2] Starch, Daniel. "The Inheritance of Abilities in School Subjects," *School and Society*, II (1915), 608-10; "The Similarity of Brothers and Sisters in Mental Traits," *Psychological Review*, XXIV (May, 1917), 235-238.

[3] Galton, F. *Inquiries into Human Faculties.* (New York: E. P. Dutton and Company, 1883) p. 157.

[4] Lauterbach, C. E. "Studies in Twin Resemblances," *Genetics*, X (November, 1925), 525-68.

pointed out that similarity in appearance as subjectively measured by Galton and similarity in quality as measured by a scale are probably not the same thing.

When the handwriting of twins was compared in this study with that of siblings, markedly higher correlations were found for twins, especially the like-sex groups. This is the finding most commonly reported in studies of siblings and twins. It is frequently accepted as lending confirmation to two well-known theories, namely, that twins are of two zygotic types, one-egg and two-egg; and that heredity is more potent in determining individual differences than environment or training.

Observation leads to the belief that the more alike twins are in mental and physical traits, the more nearly identical are their environments; the greater their dissimilarities, the greater are the differences in environment. Sex alone will make a marked difference in the milieu of twins, as will physical appearance, temperament, aptitude, and so forth. The fact that the environment of a pair of twins is not necessarily identical for both has not heretofore been sufficiently recognized in previous studies. Older twins do not show significantly greater similarities than younger twins according to correlations for younger and older twins, given in Table II. This, also, is in conformity with the usual findings in other traits and indicates that training is impotent to overcome innate differences.

In this study and others no distinct age interval exists between the younger and the older twins. To test adequately the theory that older twins show no greater degree of similarity than younger twins, sufficient data are needed for a comparison of distinct age levels, that is, the six-year-olds with the ten-year-olds or the fourteen-year-olds, and so forth. Until such data are available, the commonly accepted findings must be interpreted with caution.

The problem of symmetry reversal is an interesting one for the reason that some authorities, notably Newman,[5] maintain that mirror-imaging is an evidence of the one-egg origin of twins. One-egg twins are always of the same sex and correlations in traits are usually considerably higher for them than for the unlike-sex twins. Lauterbach has pointed out that unlike-sex twins manifest as much symmetry reversal as like-sex twins and that groups of twins showing symmetry reversal show no greater

[5] Newman, H. H. *The Biology of Twins.* (Chicago: University of Chicago Press, 1917) p. 17.

similarities in either physical or mental traits than groups which do not show symmetry reversal.[8]

As far as handwriting is concerned, symmetry reversal contributes nothing to greater similarity, according to this study. The coefficients of correlation, as given in Table II, for the like-sex group showing symmetry reversal in quality and rate are .65 and .82, respectively, and for the unlike-sex group manifesting symmetry reversal, .41 and .31, respectively. These figures do not deviate significantly from those for the other groups.

Several conclusions may be made on the basis of data presented in this study. These are as follows: (1) Twins show a greater degree of resemblance than siblings. (2) Like-sex twins show a greater degree of resemblance than unlike-sex twins. (3) Resemblance in older twins is no greater than that in younger twins. (4) Symmetry reversal is not a factor making for greater resemblance.

---

[8] Lauterbach, *op. cit.*

EDITORIALS

AN APPRECIATION

Some nine years ago a man, a thinker, an educationist, a research worker, a director of a bureau of educational research conceived the idea that there should be a journal devoted to the interests of educational research. Scanning the field of periodical literature, he found journals devoted to special subjects, as English, mathematics and history; to the interests of special groups, as school administrators, high school teachers and principals, and elementary teachers; but no journal which was making a special effort to advance education by reporting the latest techniques and results of research.

He conferred with his dean and his president and found them more than sympathetic toward the launching of such a venture. They encouraged him to do so but could give only encouragement since university funds were not available to finance the undertaking. The idea was not for a house organ or institutional magazine, but for a magazine national in scope, the editor of which would be aided by a group of associate editors, each preëminent in his field.

At this point, Mr. George A. Brown, president of the Public School Publishing Company, offered to publish such a journal, underwriting its probable deficit during its infancy, because of his interest in scientific education.

The offer was accepted and the first number of the *Journal of Educational Research* appeared under the date of January, 1920.

It is quite unnecessary to inform those who have read this journal from the start that the man who conceived the idea was Dr. B. R. Buckingham. He assumed editorial responsibility at the beginning and secured for his board of associate editors Leonard P. Ayres, W. W. Charters, S. A. Courtis, Walter S. Monroe, George D. Strayer and Lewis M. Terman. The writer who was secretary of the National Association of Directors of Educational Research, (now the American Educational Research Association) was also ex officio an associate editor due to the fact that this association had adopted the Journal as its official organ.

A statement of plan for the Journal was presented by the editor in the first issue. The extent to which the plan has been fulfilled during

the eight and one-half years Dr. Buckingham edited the Journal can be gathered by anyone who wishes to scan the nearly seven thousand pages of these seventeen volumes. Dr. Buckingham was ever alert for the best reports from this field. He not only welcomed but constantly solicited articles from the foremost research workers, yet never rejected an article containing a truly worthwhile idea because its author was unknown. He courted constructive criticism from the membership of the Educational Research Association for the betterment of the Journal.

During the last two years, pressure of other work forced him to give less [illegible] editing of the Journal and no little ere oth sea ma co in th se iss on pe la of bu ar m th ec cc of ti a I

TIGH BOU

has given [illegible]

In every capacity I regret his resignation and believe I express the feelings of those who have followed the Journal through the years that we owe Dr. B. R. Buckingham a debt of gratitude and appreciation which it is impossible to repay.

E. J. A.

REVIEWS

Address all communications relative to book reviews to Editor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

SMITH, MEREDITH. *Education and the Integration of Behavior.* New York, Teachers College, Columbia University, 1927. v + 93 pp. (Teachers College,

recent physiological investigations have been examined and from a number of points of view, it has been shown how physiological-biological material, espe-

[redacted]

Dewey's "philosophic analysis of experience." "To this end conclusions from

Educationally, as well as psychologically, this thesis means that we must

the eight and one-half years Dr. Buckingham edited the Journal can be gathered by anyone who wishes to scan the nearly seven thousand pages of these seventeen volumes. Dr. Buckingham was ever alert for the best reports from this field. He not only welcomed but constantly solicited articles from the foremost research workers, yet never rejected an article containing a truly worthwhile idea because its author was unknown. He courted constructive criticism from the membership of the Educational Research Association for the betterment of the Journal.

During the last two years, pressure of other work forced him to give less and less of his own time to the editing of the Journal and no little credit and praise must be given to Miss Josephine MacLatchy and the other members of the editorial division of the Bureau of Educational Research at Ohio State University for the high standard of editorial work maintained during this period. While he did less of the laborious work of copy-editing at this time, questions of policy and materials to be included in the Journal still received his most interested attention. Finally, feeling that his other duties demanded all his time, Dr. Buckingham, last October sent in his resignation as editor-in-chief, to be effective with the May issue.

The seventeen volumes of the Journal covering a period of eight and one-half years stand as a monument to Dr. Buckingham, who stamped his personality upon each of them. The educational world has profited largely by his carrying out this idea. Those who have had the privilege of working with him on this project have made their contribution also, but each will readily acknowledge that the chief honors are his.

As an individual who has been associated with Dr. Buckingham in an editorial capacity during the entire life of the Journal and has seen more intimately than any other the time and thought he lavished upon the Journal, I wish to express my appreciation of his services to scientific education. As one who was privileged to be a member of the executive committee of the American Educational Research Association during six of these years and consequently knew his efforts to serve this organization, I wish to pay tribute to his services. As a reader of the Journal and one of the large body of research workers in the field of education, I also wish to express appreciation for the aid and inspiration the Journal has given under his able leadership.

In every capacity I regret his resignation and believe I express the feelings of those who have followed the Journal through the years that we owe Dr. B. R. Buckingham a debt of gratitude and appreciation which it is impossible to repay.

E. J. A.

REVIEWS

Address all communications relative to book reviews to Editor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

SMITH, MEREDITH. *Education and the Integration of Behavior.* New York, Teachers College, Columbia University, 1927. v + 93 pp. (Teachers College, Columbia University, Contributions to Education, No. 261)

Some thirty-odd years ago John Dewey pointed out the limitations of the reflex-arc concept. Throughout the passage of years, however, his warning has been quite ignored and the stimulus-response formula has been elevated from its position as a descriptive device to a rôle—that of a basic psychological concept—which it has played enthusiastically despite its limitations. More recently the significance of Dewey's critique has been realized and numerous attacks upon the S-R bond have appeared. Education seems destined to discover that one of its favorite tools possesses doubtful value.

The study reported in the monograph here considered likewise shows the limitations of the stimulus-response concept. This is not its sole purpose, but, in as far as physiological and biological evidence is shown to support the philosophical and psychological interpretation of behavior as an integrated act, it lends immediate and direct support to Dewey's early contention that the reaction is *into* the stimulus rather than *to* the stimulus. The author starts out quite frankly to locate data "from experimental physiological-biological conclusions" to support Dewey's "philosophic analysis of experience." "To this end conclusions from recent physiological investigations have been examined and from a number of points of view, it has been shown how physiological-biological material, especially that of Sherrington, Child, and Herrick, supports the conception of behavior as an integrative process of organism and environment" (p. 70).

Dr. Smith shows wherein familiar psychological concepts emphasize one aspect of the total integrative process to the neglect of others equally significant. Against these approaches the author pits "consideration of the organism as a whole." This leads him to put forward a "psychological formula," *N-SR-C*, which he explains as follows:

> The *N* indicating the needs, wants or purposes which represent an integration of organism and environment, the *SR* indicating the interdependent sensori-motor action on the environment essential to meeting the needs of the organism, and the *C* referring to the environmental change produced by this action which is of consequence to the organism (p. 47).

The purpose of the formula, of course, is to direct our attention to the fact of integration. It hardly does more, however, than suggest this fact. This may be sufficient, but the chief emphasis in the formulation seems still to be placed upon the individual. If we accept integration as a fundamental concept, then "environmental need or satisfaction" becomes quite as important as the need or satisfaction of the organism.

Educationally, as well as psychologically, this thesis means that we must

the eight and one-half years Dr. Buckingham edited the Journal can be gathered by anyone who wishes to scan the nearly seven thousand pages of these seventeen volumes. Dr. Buckingham was ever alert for the best reports from this field. He not only welcomed but constantly solicited articles from the foremost research workers, yet never rejected an article containing a truly worthwhile idea because its author was unknown. He courted constructive criticism from the membership of the Educational Research Association for the betterment of the Journal.

During the last two years, pressure of other work forced him to give less and less of his own time to the editing of the Journal and no little credit and praise must be given to Miss Josephine MacLatchy and the other members of the editorial division of the Bureau of Educational Research at Ohio State University for the high standard of editorial work maintained during this period. While he did less of the laborious work of copy-editing at this time, questions of policy and materials to be included in the Journal still received his most interested attention. Finally, feeling that his other duties demanded all his time, Dr. Buckingham, last October sent in his resignation as editor-in-chief, to be effective with the May issue.

The seventeen volumes of the Journal covering a period of eight and one-half years stand as a monument to Dr. Buckingham, who stamped his personality upon each of them. The educational world has profited largely by his carrying out this idea. Those who have had the privilege of working with him on this project have made their contribution also, but each will readily acknowledge that the chief honors are his.

As an individual who has been associated with Dr. Buckingham in an editorial capacity during the entire life of the Journal and has seen more intimately than any other the time and thought he lavished upon the Journal, I wish to express my appreciation of his services to scientific education. As one who was privileged to be a member of the executive committee of the American Educational Research Association during six of these years and consequently knew his efforts to serve this organization, I wish to pay tribute to his services. As a reader of the Journal and one of the large body of research workers in the field of education, I also wish to express appreciation for the aid and inspiration the Journal has given under his able leadership.

In every capacity I regret his resignation and believe I express the feelings of those who have followed the Journal through the years that we owe Dr. B. R. Buckingham a debt of gratitude and appreciation which it is impossible to repay.

E. J. A.

REVIEWS

Address all communications relative to book reviews to Editor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

SMITH, MEREDITH. *Education and the Integration of Behavior.* New York, Teachers College, Columbia University, 1927. v + 93 pp. (Teachers College, Columbia University, Contributions to Education, No. 261)

Some thirty-odd years ago John Dewey pointed out the limitations of the reflex-arc concept. Throughout the passage of years, however, his warning has been quite ignored and the stimulus-response formula has been elevated from its position as a descriptive device to a rôle—that of a basic psychological concept—which it has played enthusiastically despite its limitations. More recently the significance of Dewey's critique has been realized and numerous attacks upon the S-R bond have appeared. Education seems destined to discover that one of its favorite tools possesses doubtful value.

The study reported in the monograph here considered likewise shows the limitations of the stimulus-response concept. This is not its sole purpose, but, in as far as physiological and biological evidence is shown to support the philosophical and psychological interpretation of behavior as an integrated act, it lends immediate and direct support to Dewey's early contention that the reaction is *into* the stimulus rather than *to* the stimulus. The author starts out quite frankly to locate data "from experimental physiological-biological conclusions" to support Dewey's "philosophic analysis of experience." "To this end conclusions from recent physiological investigations have been examined and from a number of points of view, it has been shown how physiological-biological material, especially that of Sherrington, Child, and Herrick, supports the conception of behavior as an integrative process of organism and environment" (p. 70).

Dr. Smith shows wherein familiar psychological concepts emphasize one aspect of the total integrative process to the neglect of others equally significant. Against these approaches the author pits "consideration of the organism as a whole." This leads him to put forward a "psychological formula," *N-SR-C*, which he explains as follows:

> The *N* indicating the needs, wants or purposes which represent an integration of organism and environment, the *SR* indicating the interdependent sensori-motor action on the environment essential to meeting the needs of the organism, and the *C* referring to the environmental change produced by this action which is of consequence to the organism (p. 47).

The purpose of the formula, of course, is to direct our attention to the fact of integration. It hardly does more, however, than suggest this fact. This may be sufficient, but the chief emphasis in the formulation seems still to be placed upon the individual. If we accept integration as a fundamental concept, then "environmental need or satisfaction" becomes quite as important as the need or satisfaction of the organism.

Educationally, as well as psychologically, this thesis means that we must

strike our tents. We cannot, following the lead of the external stimulus, place sole direction in the hands of the teacher, nor, after those who follow inner urges, allow the child to select his educative material as a free and unhampered agent.

From both of these opposed points of view children are responding to an environment too narrow and restricted to permit release of capacity, to give opportunity for initiative and originality in any measure of children's ability. In the one self-expression is limited to the particular form the teacher indicates; in the other it is limited to forms suggested by an environment not intentionally educative. In neither is the child really free, for freedom is not mere release from external constraint (p. 83).

This is apparently to say that neither drill nor self-initiated projects are to control in the classroom. It is true that the author's educational illustration is referred to as a "project," but in this the teacher is called upon to provide "a certain unity, permanence, or continuity in the work of the group."

Coöperation, socialization, sharing, participation are the values which this community project attempts to realize along with the knowledge needed to construct and reconstruct the environment with which the school is concerned. The technique thus illustrated is one designed to make the concept of integration illuminate the classroom practice. This reviewer feels that the important contribution of this work is the foundation presented for this view of experience—a foundation that takes us up to, or back to, Dewey's too-much-neglected thesis. This is done with considerable care. The educational implications need still to be developed, though definite "hunches" are suggested here.

H. Gordon Hullfish
Ohio State University

Holzinger, Karl J. *Statistical Methods for Students in Education.* Boston, Ginn and Company, 1928. viii + 372 pp.

Books which undertake to apply statistical methods to the solution of problems in education are becoming numerous. The book under review is the sixth textbook in educational statistics which has been published during the last three years. These publications are the natural result of the increased demand for training in statistics in schools of education, particularly graduate schools, and of the general dissatisfaction with the older, more difficult texts.

Holzinger's book, however, is not just another book on educational statistics. The reviewer does not hesitate to say that it is the most thorough, most scholarly work in educational statistics which has yet appeared. It differs from other books in its field in two important respects: first, in the detailed and thoroughgoing analytic methods presented; and second, in extent to which formulas and procedures are rationalized.

The author states in his Preface that "proofs of difficult theorems have been omitted throughout and demonstrations have been included only when experience has shown that they come within the grasp of the ordinary student." The reviewer assumes that "ordinary" means average. If this assumption is correct, Professor Holzinger is able to do what he (the reviewer) is unable to accomplish. The reviewer believes much of the algebraic manipulation to be beyond average graduate students in education. For example, the chapter on logarithms gives a brief account of arithmetical and geometrical progressions, the invention of logarithms, and the laws of exponents. This is an excellent summary for the student who has studied these topics; but because of its brevity and the lack of op-

portunity for practice on the details involved it would seem to be difficult for one of more limited mathematical training.

Another example of difficult material is found in the discussion of the mean deviation. The development here is rather unique but probably incomprehensible to many graduate students in education. In the reviewer's opinion simplifications could have been effected in both the symbolism employed and in the development of the topic.

The book is quite complete. Yet, as the author admits, many items which might have been included have been omitted. It seems a bit strange, however, that the derivation of the formula for the regression equation through the use of the calculus should be included, while the tangent concept of the correlation coefficient and the algebraic derivation of the partial-correlation formula for three variables through the correlation of residuals should be omitted. We should not be severe with the author on this point, for it is probable that no two authors would agree as to what such a book should contain.

The book is splendidly printed and the proofreading has been well done. It is sound and scholarly. It will make an excellent text for a second course if the course is elective. Gifted students having a thorough knowledge of elementary algebra—including a course in college algebra—will use the book profitably as a text for a first course, but the less gifted and the less well-trained should begin with something easier. Professor Holzinger has made a valuable addition to our rapidly growing, but not yet too large, library on educational statistics.

R. L. MORTON

Ohio University

BOOK, WILLIAM F. *How to Succeed in College*. Baltimore, Warwick & York, 1927. 192 pp.

This alluring title would lead one to suppose that the book had been written to help college students. Such, however, is not the case. The book reports for college teachers certain investigations conducted at Indiana University to reveal how students spend their time, what the conditions for study are which surround students, how well students can read, how students are distributed when rated by psychological tests, and what marks students are given by their teachers. Many important relationships are found among these several groups of data. Finally, the results of a course on "How to Study" are recorded.

These are interesting and valuable topics. The general significance of the data which are revealed will no doubt be granted although the book is not written in such a scientific way as to satisfy a critical reader. Few details of the procedures used in the investigations are given, and few precise statements are made showing how the data published were derived, and their accuracy assured.

This uncritical method of presentation takes on more seriousness when one tries to overlook the unpardonable sloppiness of the English used throughout the book. Carelessness extends beyond ambiguity and becomes definitely misleading at times. Still worse, quite indefensible statistical procedures are resorted to. On page 102 is a "ratio of efficiency" obtained by "dividing the percentage of efficiency" by the "total amount of time spent," while on page 167, with naïve disregard of the zero point, is the statement that one student "was 7.2 times as efficient in his reading" as another student.

F. J. KELLY

University of Idaho

HALLE, RITA S. *Which College?* New York, Macmillan Company, 1928. xx + 268 pp.

In this day when all young America is said to be heading for college, a number of questions loom up in regard to the choice of the institution in which to spend from four to six years. Directors of student personnel problems in the high school, the vocational- and educational-guidance counselors, and the deans of girls and boys have accepted the challenge of such questions. This little book will serve as a guide especially to those just getting their moorings in this new field of work.

There is a good deal of sound social and educational philosophy in the four introductory chapters, dealing in turn with (1) "Which college?" (2) "How to get there?" that is, by means of what entrance requirements, (3) "The junior college?" and (4) "After college,—what?" The fourth is probably the most valuable. It deals with the question, What is the legitimate function of a college education? Many are the boys and girls who have answered the question, "After high school, what?" with a matter of fact, "Why, college?" without a single sane notion as to why. This chapter attempts to put social purpose into the choice. Long before the student leaves high school he should realize that the only justification of social indulgence toward more education lies in the hope that those thus specially favored will contribute a rich return. A large measure of emphasis is placed on the responsibility involved in the selection of some adequate type of social service as a basis for making this compensation.

The remaining four chapters of the book are devoted to a discussion of the essential facts pertaining to the more than three hundred colleges and universities found on the lists of the Association of American Universities and the several regional associations under the captions, colleges for men, colleges for women, co-educational institutions, and Negro colleges. Authoritative information is given in regard to the entrance requirements, endowments, the tuition charges, the faculty, enrollment, nature of educational advantages, scholarships and loan funds, opportunities for earning money while at school, the dormitories, the fraternity and sorority situation, and many other matters of importance. Such information lends itself admirably to making tentative choices which can be substantiated by reference to college catalogues, inquiry by correspondence, and personal contact with alumni.

Obviously such a book will be welcomed by superintendents, principals, and teachers as well as those special officers already mentioned. The discriminating parent, who cannot settle the question of, "Which College?" by the choice of a convenient Alma Mater, will find here a wealth of information and inspiration. Somewhere in his high-school career every student should have recourse to these facts and access to many of the catalogues mentioned.

OSCAR F. WEBER
University of Illinois

FREELAND, GEORGE EARL, ADAMS, ROXANA MORTON, AND HALL, KATHARINE HEDGES. *Teaching in the Intermediate Grades.* Boston, Houghton Mifflin Company, 1927. xviii + 308 pp.

This book is, as the subtitle states, "A Study of Curricula and Methods of Teaching in Grades Four, Five, and Six." With its appearance, the long neglected intermediate grades have come into a method book entirely their own. Leaving the primary department, the chil-

dren have usually come into this "no-man's-land" to be bombarded by textbooks offered as a substitute for doing things which are vital and interesting. The authors emphasize pupil participation in classroom activities.

One chapter of twenty-five pages entitled "The Intermediate-Grades Child," treats original and instinctive tendencies, physical development, mental development, mental abilities not to be expected, and import of the data. This is the red-letter chapter of the book, and one only wishes there were more of it. Other chapters deal with the problems of curricula and method in reading, English, health education, geography, history, nature study and science, education for citizenship, arithmetic, writing and spelling. A chapter on dramatization is almost entirely devoted to three illustrative dramatizations prepared by pupils. Explanation of the process by which these plays evolved is lacking. The last chapter gives in detail three illustrative projects worked out by classes.

The use of illustrations followed in the last two chapters is characteristic of the book. Throughout there is an abundance of definite suggestions as to what to do, with ample reasons for so doing. There is enough illustration to stimulate the constructive imagination of the capable teacher. The book has the flavor of fresh and intimate contact with superior classroom work.

The text is adapted to class use with students. Each chapter concludes with a summary of suggestions—which should lead to participation of the readers—questions to emphasize, major issues, problems for discussion, and suggested readings. To one who has kept pace with the best in method and in practice the book has little to offer. For the host of teachers, however, who are in the intermediate grades not because they are specialists in that field, but merely because they are not teaching elsewhere, this book must be a godsend. Too, it should be exceedingly valuable to the principal who has not made special study of the intermediate-grades child and his needs.

MYRTLE L. KAUFMANN
Public Schools, Spokane, Washington

HARAP, HENRY. *The Technique of Curriculum Making.* New York, Macmillan Company, 1928. xi + 315 pp.

Considerable thought has been given to curriculum-making during the past few years. The problem has been attacked from various angles and for that reason we have several approaches to it. First, there is the activity-analysis approach. This plan is regarded by many as the most scientific, because it is definite, positive, and comprehensive. It consists of an analysis of a given task to the *n*th degree. Second, there is the objective approach which presupposes definite aims. The formulation of these objectives constitutes the primary issue in this type of curriculum-making. Third, there is the frontier thinker's approach. This approach constructs curriculum in terms of trends in civilization as predicted by original thinkers, care is taken to differentiate between basic trends and superficial changes. Fourth, there is the child-activity approach in which growth is the desired end and all subject-matter is subordinated to this end. The child under the guidance of the teacher is the ultimate curriculum-maker. The teacher's part is to stimulate and guide the child activity that growth may best ensue. Fifth, is the adult-needs approach. It selects the adult needs as the basis of curriculum-making. Sixth, there is the median present practice approach in

which the curriculum is founded on certain specific subjects determined on the basic of their survival in the changing conditions of civilization. Seventh, there is the social-statistics approach. This procedure formulates a list of objectives in a specific field by organizing objectives into teaching units and by planning activities necessary to achieve them. Eighth, and last, there is the educational-shortage approach which utilized the opinions of thoughtful educated people outside the school in determining the deficiencies of adult needs which should form the basis of the curriculum.

There is a genuine need of a treatise on curriculum-making dealing with these various approaches. Such a book would be of great value in the orientation of students of curriculum-making. This book sets forth quite clearly approaches and suggests a definite procedure in undertaking the problem. The plan presented is a result of actual experience in the field and in the classroom. The bibliographies are arranged by subjects and are perhaps the most complete ever offered on this subject. The book covers practice and theory with equal fullness. Professor Harap does not aim at a patented series of courses but rather a procedure whereby the curriculum may be adjusted to the changing activities of community life. To such an end this book is an important contribution and merits wide use. It is well written and is extremely interesting. It should be of real value to teachers, supervisors, administrators who wish help in curriculum-making.

ELLSWORTH COLLINGS
University of Oklahoma

CORYELL, NANCY GILLMORE. ***An Evaluation of Extensive and Intensive Teaching of Literature.*** New York, Bureau of Publications, Teachers College, Columbia University, 1927. vii + 201 pp. (Teachers College, Columbia University, Contributions to Education, No. 275)

Dr. Coryell's study is a complete and specific picture of carefully guarded research into a problem of importance in the teaching of senior high-school literature. By means of parallel classes of equal ability, taught by the same teacher, and so far as possible under precisely equivalent conditions, an attempt was made to find out the effects of teaching, on the one hand, a limited list of classics with the usual full and intensive consideration of detail, and, on the other hand, in addition to the same classics less exhaustively studied, to secure a great deal of varied reading of literature. The pupils in each division of the experiment were grouped on X, Y, and Z levels on the basis of previous ratings in English, Terman intelligent quotient, and reading age and reading quotient (Thorndike-McCall).

The tests used at the end of the period of instruction were chiefly objective tests of memory of what was read, although attention to comprehension, and in some degree to appreciation, was also given. The reliability of these tests, which were carefully checked, ranged from +.64 to +.70; on all tests taken together, the coefficient was +.90. The uniform school examinations of the Board of Regents, Thorndike's Test of Word Knowledge, and the Abbott-Trabue and Logasa-McCoy exercises in appreciation were all used. In addition to at least an equal degree of comprehension and appreciation of the literature studied, the extensive-reading group covered six times as much literature as the intensive group. The greatest improvement in reading ability occurred in the extensive-

reading group which had the lowest initial scores; these pupils made an "average gain of 17 points, or three times the normal gain between the eleventh and twelfth year."

This study shows the great value of cordial coöperation between Departments of English and Departments of Education in the care with which the experiment was checked and guarded, both in relation to the statistical procedures involved and to the knowledge of the subject-matter concerned and of all the important literature on that subject. The monograph deserves the careful and full consideration of everyone concerned with building high-school courses in literature. In particular, the stenographic reports of every type of lesson used, with the interesting evaluation of the procedures involved, will give a great deal of help to the teacher and supervisor of English.

S. A. LEONARD
University of Wisconsin

ROGERS, FREDERICK RAND. *Tests and Measurement Programs in the Redirection of Physical Education.* New York, Teachers College, Columbia University, 1927. (School Administration Series). 166 pp.

Mirabile dictu! A book on physical education in which the author gives major stress to underlying pedagogical principles; a book in which social efficiency values are stressed and a usable program set up by which such objectives may be attained; a program in which physical measurements becomes a potent factor in the selection of teams and the determination of physical activity!

Such a book and such a program is a joy to one who sees physical education as an important element in the complete education of boys and girls.

The explanation for the point of view is doubtless found in the author who is first a school man and second a physical education expert. The doctrine is well expressed in the following words: "In competitive activities, educational objectives will be achieved most effectively only when the powers of competing individuals or teams have been equalized." When this is put in practice real sportsmanship can be developed, because the ideal is changed from winning by means of a heavier or older or more experienced team to the joy of sportsmanship where the contending forces are as nearly equal as possible. Dr. Rogers has developed a series of tests and a formula for strength index and for physical fitness index which are highly reliable. The cost of the instruments for the measuring program is not excessive and the time element is well within that which can be given for such a purpose. The procedure used in the development of the tests, together with the statistical treatment for determining reliability and objectivity, are clearly set forth but the heart is in the program which provides for meeting individual needs and placing the whole procedure on a sound pedagogical basis. School superintendents who wish the greatest possible good to their pupils from the physical education program (and what one does not?) and coaches who think in terms of all the children instead of those only who need physical work least will find the volume interesting and instructive.

E. J. ASHBAUGH
Ohio State University

ROGERS, FREDERICK RAND. *Tests and Measurement Programs in the Redirection of Physical Education.* New York, Teachers College, Columbia University, 1927. (School Administration Series). 166 pp.

Dr. Rogers joins the ranks of testers in Physical Education, along with Messrs. D. K. Brace and F. W. Cozens, in a partial revival of one of the old phases of Anthropometry. This book is quite in accordance with the nation-wide craze for testing in all educational fields, and is up-to-date in its use of statistics. From the viewpoint of the physical educator, even though unversed in statistical methods, some very practical suggestions are made and the laboratory procedure of applying them is also given. To the principal, the idea of sending pupils to the gymnasium not by grades but according to physical ability, in relation to age, height and weight, is to most of them unprecedented.

The three main topics in the book are the derivation and use of: (1) The Strength Index (S. I.); (2) The Physical Fitness Index (P. F. I.); (3) An Administrative Program in Physical Education. Under the latter we are told of an "Athletic Index and a Program for Coaches." In this day of "athletic monopoly," in most physical education programs, it is worth while to consider means for equalizing competition and setting physical standards for eligibility to play on "varsity" teams.

The measurement program is not a substitute for skillful teaching but a valuable aid to the most skillful teacher. The physical educator of to-day recognizes "that measures of muscular strength bear similar relationship to Physical Education that measures of intelligence bear to mental education." For these reasons, the book is recommended especially to physical educators but also to all educators who profess to know modern trends of Education.

SAMUEL H. COBB
Ohio State University

BRANDENBURG, G. C., and REMMERS, H. H. *The Purdue Rating Scale for Instructors.* Lafayette Printing Company, Lafayette, Indiana.

To those instructors to whom a frank and systematic estimate of themselves as instructors is desirable, The Purdue Rating Scale for Instructors will be of value. This Scale is a well-organized rating scale on which ten traits presumably vital to effective classroom teaching are supplied together with an orthodox device for the judge to use in estimating the amount of each trait the instructor rated possesses. The authors are sophisticated enough to insist, in the directions, that no rater sign his name or otherwise identify himself.

A compact, well-written manual accompanies this scale. Data are supplied to show that a rating by 120 pupils will give a reliability (or validity which seems identical in this instance with reliability) of +.95. In other words if an instructor will use this scale with 120 pupils according to simple directions as printed he will find out with sufficient exactness just what he seems to be as far as the student group mind is concerned. Other statistical tables of merit relative to the behavior of the scales are printed in the manual. A use of this scale should supply to an instructor confidential information by which he may well profit in attempts to increase his acceptability to his students.

F. B. KNIGHT
State University of Iowa

RESEARCH ABSTRACTS and BIBLIOGRAPHIES

Address all communications relative to research abstracts and bibliographies to Editor A. S. Barr, University of Wisconsin, Madison, Wisconsin.

CLARK, HAROLD F. "The Economic Effects of Education" *Bulletin of the School of Education*, Bloomington, Indiana: Indiana University, 1928.

This bulletin is a collection of statements of economists of the last fifty years on the economic effects of Education. Most of the quotations express the view that more education would increase economic welfare. Any one kind of education can, however, be oversupplied from the economic standpoint. "Educated labor does not receive relatively high wages because it is educated but because there is a scarcity of educated workmen." Schooling has a tremendous power in directing people into channels other than the most productive ones. Such is the more or less inevitable result when a school system is built in one age for a small percentage of the population and is used in a later age for educating a very much larger percentage of the population. A fair number of writers intimate that more schooling would tend to equate income, thus destroying the monopoly now enjoyed by the few. While schooling in certain cases thus does not increase the income of the individual getting the education it does not follow that it is economically bad. It is entirely possible for schooling to fail to increase the earnings of the individual and at the same time to enormously increase the total income of society. The writer concludes by raising four questions:

1. What is the effect of schooling upon the income or earning power of the individual?
2. What is the effect of schooling upon the total income of society?
3. How far and under what conditions may schooling of certain types tend to lower the income of the individual and at the same same time increase the income of society?
4. What is the effect of education upon production, distribution, and consumption?

DOUGLASS, HARL R. "The Experimental Comparison of the Relative Effectiveness of Two Sequences in Supervised Study," Eugene, Oregon: *University of Oregon Publication*, 1927.

The above bulletin is an experimental study of relative effectiveness of two sequences in supervised study,—(1) the study-recite plan in which recitation is preceded by study, and (2) the recite-study plan in which the recitation precedes the supervised study period. The experiment was carried on in the University of Oregon high school with ten paired sections in English, Science, Mathematics and History with the following results:

1. It cannot be said that either an R-S or an S-R sequence is more effective than the other for all classes or types of work.
2. The S-R sequence is more effective in classes in history and social science and in literature than the R-S sequence.
3. The R-S sequence is more effective in classes in mathematics and science,

though the superiority may not be manifest in each class.

4. The relative effectiveness of the S-R sequence for history and social science and English classes is greater for classes in grades above the eighth than for the seventh- and eighth-grade classes.

5. Teachers' judgments as to relative effectiveness of two methods are not to be taken too seriously.

6. Neither sequence operates to produce greater variability in progress than the other, generally or in any particular subject or school grade.

7. Neither sequence is peculiarly favorable for the more capable or for the less capable student.

"Guidance in Secondary Schools." *Bulletin of the National Association of Secondary-School Principals.* Report of Committee on Guidance. Bulletin No. 19, January, 1928.

The major purpose of this study is to supply information of a basic sort which will enable the secondary schools to formulate and to carry out a systematic program of purposeful guidance. This bulletin was prepared with the idea of meeting particularly the needs of the smaller secondary schools.

Guidance in education has come to mean more than assistance to pupils in planning their school and vocational careers; it has been extended so as to indicate not only educational guidance but also health guidance, moral guidance, guidance for leisure, guidance for home life, and guidance for social relationships. It is with guidance in this broadened sense that this bulletin mainly has to deal.

MARBERRY, J. O. *The Administration of Public Education in Centralized and Coördinate Schools.* Austin, Texas: Bureau of School Inquiry, University of Texas, 1928.

The problem of this study was to compare the efficiency of the centralized and the coördinated types of school administration. A centralized school is one having all authority centering in the office of the superintendent of schools; a coördinated school is one having a business manager who is responsible in fiscal matters directly to the board of education and not to the superintendent of schools.

The school systems selected for this study included cities ranging in population from 30,000 to 250,000. Complete data were obtained from 206 of these cities, 140 of which followed the centralized type of school organization, and 66 followed the coördinated type of organization.

Two measures of administrative efficiency were applied to these schools: (a) an index rating and (b) a functioning rating. The index number was computed from the following educational and financial factors:

1. Teacher-pupil ratio in the elementary grades on the basis of average daily attendance.

2. Teacher-pupil ratio in the secondary grades on the basis of average daily attendance.

3. Percentage secondary teachers are of total teachers in day schools.

4. Percentage average daily attendance of day schools.

5. Average time spent per pupil in daily attendance on the basis of current expenses.

6. Percentage salaries of teachers in day schools are of current expenses.

7. Percentage supervisors and principals salaries and expenses are of current expenses.

8. Percentage general control is of current expenses.

These index numbers were subjected to statistical treatment to determine differences if any in the efficiency of centralized and coördinated types of schools. The critical ratio for each index number is given in Table I.

TABLE I

	Index Numbers							
Type of Administration	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Centralized	....	....	2.38	2.20	....	1.04	2.00	×
Coördinated	1.12	2.23	....	....	5.90	....	....	0

As a critical ratio must be at least 3 times the difference to be statistically significant, the conclusion that one must draw from these data is that there is no significant difference between the efficiency of the two types of administration except for index number 5, which is the average amount of time spent per pupil in daily attendance. This critical ratio favors the coördinate type of school administration.

A second measure of the relative effectiveness of these two types of school organization was computed by giving a value of 10 to functions which school administrators generally agree should be initiated or executed by the superintendent of schools. Ten such functions were included. Consequently a school system in which the superintendent initiated or executed all these functions would have a score of 100, and such a school system would be unexcelled in efficiency as far as expert opinion is concerned. The functions of school administration which it was assumed that the superintendent of schools should initiate were as follows:

1. Appointment of assistant superintendents, principals, teachers, janitors, and clerks.
2. The preparation of the budget.
3. Transfer of principals, teachers, and other employees.
4. Dismissal of principals, teachers, and other employees.
5. Taking the census and enforcing the compulsory attendance laws.
6. Pertaining to purchase and sale of buildings and grounds, approving architects' plans, and maintenance of buildings.
7. Determination of the curricula.
8. Selection of textbooks and supplies.
9. Determining new policies.
10. Direction and supervision of classroom instruction, continuation schools and evening schools.

The data secured from this study were submitted to statistical treatment. The critical ratios for the several items studied are given in Table II.

There are significant differences in favor of the centralized school in the first six assumed functions of the superintendent of schools. While two of the remaining differences are opposed to the centralized school plan they are not statistically significant.

TABLE II

	Functions of Administration									
Type of Administration	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Centralized	7.87	8.52	7.20	5.40	7.43	5.95	....	1.95	1.45	....
Coördinated	....	....	....	....	....	....	2.10	....	....	.40

SIMS, VERNER MARTIN. *The Measurement of Socio-Economic Status.* Bloomington, Illinois: Public School Publishing Co., 1928.

This study explains the procedure employed in the construction of a score card to measure socio-economic status.

The factors determining the socio-economic status of children were first studied. In order to secure this information 55 questions which could easily be answered by pupils were prepared. The following points were considered in preparation of these questions:

1. Each question must be indicative of the economic or the cultural level of the home.
2. The questions must cover as many aspects of the home background as possible.
3. The questions must be so stated that the child can understand them.
4. The questions must ask for information which the child can furnish.
5. The questions must ask for information which the child is willing to furnish.
6. The questions must be stated in such a manner that there will be a minimal chance of error.
7. The questions must make provisions for answers that are comparable.

When these criteria were applied to the 55 questions 8 were eliminated leaving thus 47 questions which composed the preliminary blank known as Form B. Form B was then given a preliminary trial in the sixth, seventh, and eighth grades of six public schools in New Haven, Connecticut. These schools were selected by the superintendent of schools because, in his opinion, they represented a socio-economic cross section of the school population of that city. The form was filled out by 686 pupils.

The blank was found to be cumbersome in its preliminary form and consequently the number of questions was reduced to somewhere between 20 and 25. The final number retained in the score card (Form C) was 23. The following criteria were used in determining the 23 questions to be retained in Form C:

1. The ability and willingness of the persons tested to furnish the information.
2. The correlation of each question with the total for other questions.
3. The inter-correlation between items.
4. The percentage of the population possessing the item asked for in the question.
5. The reliability of the questions.
6. The variety of aspects of home background measured.
7. Common sense.

Some of the items enumerated seemed more fundamental in determining the socio-economic status than others. Consequently, weights were assigned to each item. To secure a socio-economic status score the credits for each question are added and the sum divided by the number of questions answered. The quotient carried to one decimal place, and the decimal neglected, becomes the child's socio-economic status score. This method of scoring is justified because the results obtained are found to correlate $.95 \pm .01$ with more complicated methods of scoring.

The validity of the score card was determined by checking the socio-economic status of pupils in three schools of radically different socio-economic status. The score card seemed to satisfactorily differentiate these pupils. The following items are included in the final score card:

1. Telephone, 2. Furnace, 3. Bathroom, 4. Bank account, 5. Room-persons, 6. Servants, 7. Auto, 8. Dental work, 9. Vacations, 10. Occupation-father, 11. Father-college, 12. Mother-college, 13. Father-

high-school, 14. Mother high-school, 15. Mother-lectures, 16. Mother-clubs, 17. Concerts, 18. Books, 19. Magazines, 20. Own room, 21. Clubs, 22. Music lessons, 23. Dancing lessons.

The score card gives education an instrument by which a quantitive measure of the home environment can be made.

SOST, J. *Nature and Significance of School Certificates.* (Wesen und Bedeutung der Schulzeugnisse), *New Contributions to Education,* (Neue Beitraege zur Erziehungswissenschaft), No. 1, edited by Dr. Friedrich Schneider, Univ. Koeln, Verlag: Ferdinand Schoeningh, Paderborn, Germany, 1926.

This is a doctor's dissertation in education from the University of Cologne. Dr. Sost makes in the first part of his report a critical study of the nature of a mark, errors in assigning marks, value and significance of certificates, etc. In a second part he gives the outcomings of a questionnaire dealing with the different viewpoints and standards of assigning marks by German teachers. In the last part, Dr. Sost offers a statistical treatment of 65,123 school marks. He is especially interested in the distribution of these marks in relation to the normal curve, correlations between the different school subjects, sex differences in school performance, the relation between pupils' performances and the social classes from which they come, and repetition and its causes.

Dr. Sost found in one of his experiments that 105 teachers marked the same paper in penmanship with about 80 percent agreement in spite of the fact that the teachers were from different school systems. If these results are accepted, there would seem to be a much greater reliability among German teachers than American teachers in assigning school marks. Dr. Sost also found that school marks did not distribute themselves according to the normal, but always according to a one-peak asymmetrical curve. He thus objects to schemes of statistically assigning school marks according to the normal curve. The curve of school marks of boys had their median at fair (or 3) and of girls at good (or 2).

In the statistical treatment of his data Dr. Sost compares three formulas of correlation: namely the rank-order formula, the product moment formula and a third formula devised by Prof. G. Deuchler. In working over his data with these three formulas he discovers large differences in the correlation-coefficient when calculated by these three methods. For instance between arithmetic and reading he found .95, .60, and .42.

In criticizing the sensitiveness and relativeness of these formulas he asks for a better understanding of the coefficient of correlation. The highest correlations were found between reading, oral expression and orthography, i.e., $.709 \pm .058$; $.652 \pm .067$. The correlation coefficients between the different subjects increased considerably from the first grade to the fourth. Thus between arithmetic and reading the figures for the first four grades are .51, .71, .80, and .82.

In his chapter on sex differences Dr. Sost reports that girls on an average make in all subjects higher marks than boys, with a slight exception in arithmetic and drawing. The cause of this he attributes to a greater zeal of girls for school work.

The study of the dependence of school performance on the social status of pupils showed that the upper quarter of marks were assigned, 51 percent to children from the higher social class, 36 percent to children from the working class. The low social class also had repeaters,

namely 10 percent as against 1 percent for the high class.

It was found that most children fail in the first grade (about 12 percent), and also fail in orthography. From these data Dr. Sost concludes that the first two school years are critical years in the pupil's school career. The percentage failing is on an average about 7½ from each grade each year. Only about 48 percent of the children from the working class go straight forward through the eighth grade.

TIEGS, ERNEST WALTER. *An Evaluation of Some Techniques of Teacher Selection.* Bloomington, Illinois: Public School Publishing Co., 1928.

This study is an evaluation of some techniques of selecting teachers. The writer presents, first, an account of present practices; second, a summary of previous contributions to the study of techniques for predicting and judging success; third, an experimental evaluation of the letter of application, the photograph, and the recommendation; fourth, a study of the graphic rating technique, and fifth, an evaluation of trade and mental tests and the application blank in the selection of teachers.

The first experimental study was made of the letter of application, the recommendation, and the photograph. There were available in the office of the Minneapolis Board of Education, (the study was carried on in the Minneapolis Public Schools), in July 1925 the complete records of 100 elementary school teachers for grades one to six, inclusive, who had been appointed to begin service in September of that year. The complete records of twenty-five of these were withdrawn at random for ranking purposes. These twenty-five teachers were ranked, first, on the basis of the candidates' recommendations; then on the basis of letters of application; and finally on the basis of photographs. Each of these rankings was done by eleven judges. After ranking the candidates on the three separate elements named above, the complete record of each teacher was assembled and five judges again ranked them on the basis of desirability as a teacher. The data presented indicated that there were wide variations between the ranks assigned the same teacher by different judges, the intercorrelations between 5 judges on recommendations being —.37 to +.78, between 6 judges on letters of application being —.11 to +.62, and between 5 judges on photographs being, .00 to +.50. The coefficient of reliability between first and second rankings of recommendations (5 judges) was .00 to +.37. The reliability for the other three rankings was not given but can be assumed to be correspondingly low. The correlation between recommendations and teaching success was .08, between letters of application and teaching success —.04, between photographs and teaching success —.08; between the composite of these three ratings and teaching success —.13; and between the complete file with teaching success +.27. The criterion of teaching success was a group rating made by the principal, special supervisors, the supervisors of probationary teachers and the assistant superintendent in charge of this work. The official rank consisted of thirteen points on which each teacher received an excellent, good, fair, or poor. Allowing "1" for excellent, "2" for good, "3" for fair, and "4" for poor, the official rank was obtained by adding the scores on the thirteen items.

The author next experimented with one graphic rating scale method using both a twelve-point blank and a forty-one-

point blank. The twelve-point blank included such items as health, character, loyalty, force, coöperation, interest, tact, personality, mental power, professional knowledge, teaching type and teaching ability. Each item was carefully defined. The chief points brought out by this investigation were:

1. When 77 school officials rated 77 teachers twice (two to four weeks intervening) on the twelve-point blank, the coefficient of reliability of the blank as a whole proved to be $+.79 \pm .03$; the reliability of individual items ranged from $+.14 \pm 08$ to $+.84 \pm .02$.

2. Two special supervisors agreed with twenty-one principals in rating twenty-five teachers on the twelve-point blank to the extent indicated by a coefficient of $+.70 \pm .07$.

3. The coefficient of correlation (rank-difference method) between prediction of success (50 teachers on twelve-point blank) by school officials outside of Minneapolis and actual performance as determined by the official ratings was $-.21 \pm .09$.

4. When 24 school officials rated 75 teachers twice on the forty-one-point blank (one day to two weeks intervening), the coefficient of reliability for the blank as a whole was $+.93 \pm .01$; the reliability of individual items varied from $+.52 \pm .06$ to $+.91 \pm .02$.

5. A study of the reliabilities of twelve items which appear on both the twelve-point and forty-one-point blanks reveals the fact that six retain the same rank order or are displaced not more than one point; three are displaced not more than three points; and three are displaced five or more points.

Finally three other measures of teaching ability were tested, namely, a trade test, (Knight's Professional Tests for Elementary Teachers), an intelligence test (Miller's Mental Ability Test) and the Minneapolis Application Blank. The facts brought out by this study were that:

1. The coefficient of correlation between the trade test scores and official ratings of 80 teachers was $+.02 \pm .08$.

2. The coefficient of correlation between the intelligence quotients and official ratings of 80 teachers was $+.01 \pm .08$.

3. Intercorrelations between trade and mental test scores, prediction scores on inquiry blanks and official ratings range from $-.21 \pm .08$ for inquiry blank rank and official rank, and $+.54 \pm .06$ between scores on trade and mental tests.

4. The coefficient of correlation between rank as determined by scoring application blanks and official rank for twenty-five teachers was $+.27 \pm .12$.

EXPERIMENTAL STUDIES REPORTED IN CURRENT ISSUES OF OTHER MAGAZINES

PSYCHOLOGY OF LEARNING

BROWN, ROBERTA W. "A Comparative Study of the 'Whole,' 'Part,' and 'Combination' Methods of Learning Piano Music," *Journal of Experimental Psychology*, XI (June, 1928), 235-247.

In the whole-method of learning piano music the score was played from beginning to end without stopping to correct or repeat measures where errors occurred. In the part-method the score was divided into units and each unit was practiced an equal number of times. In the combination-method the score was played from beginning to end but all measures where errors occurred were repeated an equal number of times. Of the three methods

studied, the whole-method of learning piano music was the most efficient in two units out of three; the combination-method was most efficient in one instance; and the part-method was the least efficient of all.

WORCESTER, D. A. "The Permanence of Learning in High School Subjects," *Journal of Educational Psychology*, XIX (May, 1928), 343-345.

A first report of an extended investigation which is under way to determine the amount of specific content of various school subjects retained after varying periods of time.

EXPERIMENTAL STUDIES OF TEACHING

BOGARDUS, EMORY S. "Teaching and Social Distance," *Journal of Educational Sociology*, I (June, 1928), 595-598.

Non-experimental in character, but suggestive. According to the author efficiency in teaching varies inversely with the social distance between teacher and pupil, other things being equal. The greater the social distance between teacher and pupil, the less likely is the pupil to enter into the teacher's thinking, the less likely is the teacher to "connect" with the pupil's thinking, and the more likely is education to become unnatural and formal. The greater the distance the less likely that the pupil will want to learn from his teacher.

DUEL, H. W. "The Even-Front System Versus the Rotation System in Laboratory Physics," *The School Review*, XXXVI (June, 1928), 447-454.

An experimental study of the even-front system versus the rotation system in laboratory physics. The even-front system is a laboratory system in which all the members of a class are performing the same exercise at the same time. Under the rotation system a number of different exercises are being performed at the same time, a limited number of pupils being assigned to each set of apparatus. The study was made in the West High School, Minneapolis, during the school year 1925-26. The subjects of the investigation were 168 pupils in second-term physics. The experiment was of the equivalent-group type. The author summarized his finding as follows:

1. The rotation system yields better results than does the even-front system in imparting essential information.

2. The use of the rotation system means a saving in laboratory apparatus and equipment.

3. The rotation system is more economical of laboratory time and permits the performance of more exercises or the use of more time in recitation.

4. The rotation system is to be preferred because it saves much of the time usually devoted to the care of apparatus.

5. The rotation system puts a premium on industry, while the even-front system offers very little incentive and opportunity to industrious pupils.

HOLLAND, B. F. "The Effect of Class Size on Scholastic Achievement in Educational Psychology," *School and Society*, XXVIII (June 2, 1928), 668-670.

Four hundred ninety-two students in educational psychology were divided into eight sections as follows: 86, 112, 63, 71, 61, 42, 34, and 26. Each section was given the same assignment, the same tests, and the same final examination. From a study of the average grade made by each class the author concludes that size of the class is not an important factor in learning under the lecture method of instruction.

EDUCATIONAL AND VOCATIONAL GUIDANCE

PIERSON, CELIA D., and NETTELS, CHAS. H. "A Study of High School Seniors to Determine Who Shall Be Recommended to College," *School and Society*, XXVIII, (Aug. 18, 1928), 215-216.

An attempt by means of regression equations, using as variables academic achievements, character ratings, and intelligence quotients, to predict success in the freshman year of college.

EDUCATIONAL TESTS AND MEASUREMENTS

CHEYDLEUR, FREDERIC D. "Results and Significance of the New Type of Modern Language Tests," *The Modern Language Journal*, XII (April, 1928), 513-531.

A discussion of the results and the significance of the new type of modern language tests. The writer has selected certain findings derived from the use of the American Council tests which seem characteristic of the larger body of material from which they are drawn. After reviewing these materials he reaches the following conclusions:

1. The overlapping of school grades and classes in achievement as shown by the results of the American Council French Grammar Test and other standardized language tests calls for both a more scientific method of measuring the achievement of our students and of sectioning them into classes.

2. The study of the comparative achievement in schools reveals an astonishing difference in the level of work performed in various institutions and reveals the need of remedial measures, such as, programs of study, more forceful teaching, and the use of objective tests for placement purposes.

3. The proper articulation of high-school and college courses, because of varying standards of work, cannot be said to be generally effective under the present unit system at accrediting schools.

4. The effect of method on achievement as brought forth by these data would indicate the advisability of reform among direct methodists.

5. The influence of Latin training on French seems to bring encouraging evidence of the transfer power or carry-over of one subject to another.

6. And finally the comparison of functional and theoretical grammar, seems to bring forth data in favor of the former.

INTELLIGENCE TESTING

ADAMS, HENRY F., FURNISS, LOUIS, AND DEBOW, L. A. "Personality as Revealed by Mental Test Scores and by School Grades," *The Journal of Applied Psychology*, XII (June, 1928), 261-277.

A study of methods of admission to universities, or mental makeup as revealed by mental tests and school grades. After reviewing studies by Pressey, Brown, Laird, Sangren, and Webb and after presenting their own data based upon correlations between a composite of 20 rankings upon 63 personal characteristics with school grades and mental test scores for 96 subjects, the authors conclude:

1. That school grades are a better measure of mentality than are mental tests.

2. That school grades are a better measure than intelligence tests of foresight.

3. That school grades are better selective instruments than mental tests for determining the reliability and dependability of individuals.

4. That school grades are better instruments for selecting the resolute or

determined sort of person than are mental tests.

5. That school grades are measures of both mentality and character.

AUTENRIETH, WALTER. "Evaluation of the Total Personality of a Pupil." (Zur Bewertung der Gesamtpersonlichkeit des Schulers). *Zeitschrift für padagogische Psychologie, experimentelle Padagogik und jugendkundliche Forschung.* XXIX (Feb. 1928), pp. 65-79.

The problem of this investigation was to combine already analytically established single traits into one total measure which is usable for comparison of one individuality with another. The author ranked 28 pupils according to a personality scale (PS) of the professional guidance bureau of the city of Frankfort am Main. The scale contains the following items: (I) Strong abilities or disabilities, inclinations or disinclinations; (II) Perceptual types (visual, acoustic, motor, and mixed); (III) Character: subjective, objective and such special traits as: (1) Power of comprehension, (2) Memory, a) mechanical, b) logical, (3) Ability to think, (4) Imagination, (5) Enthusiasm, (6) Willingness to sacrifice, (7) Social adaptability, (8) Honesty, (9) Resoluteness, (10) Independence, (11) Leadership, (12) General adaptability (Anstelligkeit), (13) Self dependence in technique of learning, (14) Speed of learning, (15) Endurance, (16) Reliability, (17) Love for order and system, (18) Conventional behavior.

The above mentioned pupils were observed for more than 3 years. For validating this scale Mr. Autenrieth used three independent outside criteria as follows: (1) teacher's estimates of the personality value (similar to estimates of intelligence) which were made by two independent teachers. (PEst) The correlation coefficient between these two teachers' estimates is .89 ± .03. (2) school performances of the pupils, i.e., taken from school marks (M). (3) a complete mental examination and professional aptitude tests which were not taken by all of the pupils. The rank order correlation coefficients between PS and PEst' are .88 ± .03; and .73 ± .06; between PS and M .74 ± .07; between PEst' and M .90 — .02. The results of the mental examination, so far as determined, lie between the results of PS and PEst'— PEst." The contribution of the author is in the fact that scores of single analyzed abilities (in the PS scale) are unified to one total measure of the personality which stands in a meaningful correlation to the values gained by estimates and other performances of the personality.

HUNT, THELMA. "The Measurement of Social Intelligence," *The Journal of Applied Psychology,* XII (June, 1928), 317-334.

A study of the George Washington Social Intelligence Test, its validity, reliability, and relation to abstract intelligence, mechanical intelligence, age, sex, and scholarship. The test is composed of six parts: (1) Judgment in Social Situations, (2) Memory for Names and Faces, (3) Recognition of Mental States from Facial Expression, (4) Observation of Human Behavior, (5) Social Information, and (6) Recognition of the Mental States Behind Words. The norms are based upon some 6,000 students in educational institutions and some 1,000 individuals in various industrial groups.

JOHNSON, H. M. "Some Fallacies Underlying the Use of Psychological 'Tests,'" *Psychological Review,* XXXV (July, 1928), 328-337.

A 'test' is an indirect means of measuring an empirical and variable quantity,

which we may call Y, by measuring another variable, X, upon whose values the corresponding values of Y depend. The conditions which must be fulfilled before any X-variable can be used to test any Y-variable are: (1) both Y and X must be capable of being measured directly and independently; (2) the values of Y must depend on those of X in a manner which can be specified. If Y is not amenable to direct observation and measurement, can its existence then be assumed? Are not 'intelligence,' 'learning-capacity,' 'arithmetic ability,' 'interest' and 'fatigue' variables which cannot be observed and measured but for which psychologists have been trying to find 'tests'?

LINCOLN, EDWARD A. "Studies of the Validity of the Dearborn General Intelligence Examinations," *Journal of Educational Psychology*, XIX (May, 1928), 346-349.

A brief report of some validity studies made of the Dearborn Series I Tests in their original form. Five groups of children were given Dearborn General Examinations 1, 2, and 3 and the Stanford Revision of the Binet Tests. The correlations in groups examined by experts were very satisfactory (.83 to .95). The lower correlations (.66 to .71) for groups not examined by experts emphasizes the fact that trained examiners are needed for both group and individual intelligence testing.

OATES, DAVID W. "An Experimental Study of Temperament," *The British Journal of Psychology*, XIX (July, 1928), 1-30.

A critical investigation of the June Downey Will-Temperament Test. The author presents in addition to his own data which are (a) marks obtained in the Terminal Examination, (b) scores from a battery of ten intelligence tests, and (c) score from the Downey Will-Temperament Test, for fifty secondary-school boys, a review of data presented in earlier studies by Meier, Ruch, Bryant, Herskowitz, Collins, and Downey. His conclusions are:

1. A study of selected profiles indicates that the tests have some practical value in so far as they apparently enable us to predict the type of reactions that individuals will make who possess these qualities to a marked degree. They have decidedly less differential value in normal cases where there is no very definite temperamental tendency.

2. The marking of some of the tests is subjective and a modification of the scoring of some of the tests is necessary.

3. The factor or factors causing the correlation between intelligence and scholastic ability do not enter into temperament. Scholastic ability depends upon intelligence and temperament which are apparently independent of one another. The partial correlation of intelligence with temperament with scholastic ability held constant is .0006.

4. Four Will-Temperament Tests appear to have value in predicting success or failure in school studies. The group with highest scholastic success were superior in coördination of impulses and motor inhibition; the group with the lowest scholastic record were superior in self-confidence and finality of judgment.

5. The Spearman criterion for the operation of a general factor was not satisfied, but there appeared to be a group factor of wide generality.

6. While the tests may not be measures of the traits which their names suggest, they represent a technique and a line of attack which must play a prominent part in the future investigation of human personality.

SLOCOMBE, C. S. "Three Points of Interest to Mental Constructors," *The British Journal of Psychology*, XIX (July, 1918), 31-33.

It was thought that possibly there would be a difference in the ability of pupils to perform mental operations which might be termed Inventives (recall) and Selective (multiple-choices). Each of seven sub-tests, analogies, opposites, classification, effect relations, synonyms, narrative completion, and reasons, were set in two forms. The tests were tried out with groups of children ages 11-12. The author concludes from data presented that:

1. Tests set in the selective form and in the inventive form measure the same type of intelligence.

2. The selective form is the more accurate measure of that intelligence.

3. Synonyms and narrative completion are the best types of sub-tests for children aged 11-12.

NEW TESTS

An Achievement Test in Solid Geometry. Louis A. McCoy. Reported in *Mathematics Teacher.* (March, 1928), 151-162.

New York Latin Achievement Test. Harold G. Thompson and Jacob S. Orleans. World Book Company, 1928. Test I, Forms A and B, for the first half year; Test II, Forms A and B for the second half year.

Sangren-Woody Reading Test. Paul V. Sangren and Clifford Woody. World Book Company, 1928. Forms A and B.

Denny-Nelson American History Test. E. C. Denny and M. J. Nelson. World Book Company, 1928. Forms A and B.

Otis Self-Administering Tests of Mental Ability. Arthur S. Otis. World Book Company, 1928. Form C.

Public School Achievement Tests. Jacob S. Orleans. Public School Publishing Company, 1928. Tests in ten elementary school subjects. Battery A booklet is composed of tests in (a) Reading, (b) Arithmetic Computation, (c) Arithmetic Reasoning, (d) Language Usage, and (e) Spelling; Battery B booklet is composed of tests in (f) Grammar, (g) History, and (h) Geography; and Battery C booklet is composed of tests in (i) Nature Study and (j) Health. The test in each subject is also published separately.

Achievement Test in Plane Geometry. Maude McMindes. Public School Publishing Company, 1928. Form A for first testing; Form B for second testing.

Test in Civic Action. Howard C. Hill and Howard E. Wilson. Public School Publishing Company, 1928.

Précis Test. Irvin C. Poley. Public School Publishing Company, 1928.

Diagnostic Tests in the Fundamental Reading Habits. Luella C. Pressey. Public School Publishing Company, 1928. Tests for grade 2, for grades 3 and 4, and for grades 5 and 6. Forms A and B of each test.

English Vocabulary Tests for High School and College Students. W. T. Markham. Public School Publishing Company, 1928. Form 1 for first Testing; Form 2 for second testing.

NEWS and COMMUNICATIONS

Address all research news and communications to Editor Earl Hudelson, College of Education, University of Minnesota, Minneapolis, Minnesota.

At the March, 1927, meeting of the National Association of Secondary School Principals, held in St. Louis, a Committee was created for the purpose of promoting coöperative investigations in the field of secondary education. The Committee is now undertaking the organization of a number of centers of discussion and investigation. The plan is to stimulate high-school principals in some forty or fifty parts of the country to organize themselves into clubs which will meet once a month during each school year and conduct coöperative inquiries into some of the problems which are of general interest to all high-school officers. Problems suggested for early attack are: large and small classes; high-school records; methods of judging the merits of teachers; and problems of curriculum revision or construction and of financial administration. Under the chairmanship of Dr. C. H. Judd of the University of Chicago, the Committee is assembling reference material and will act as a clearance house for the various studies undertaken.

Under the auspices of the National Council of Teachers of English, Professor S. A. Leonard of the University of Wisconsin recently made a six-weeks tour of the research centers over most of the country to ascertain what research studies helpful to teachers of English are in process or completed but unpublished. Dr. Leonard reports a great and increasing interest in research in this large and difficult field. Evidence of his impression is the formidable array of studies listed in his mimeographed report. Evidently the popular belief that English teachers are indifferent toward research is based in part at least upon ignorance of what is actually being attempted. Dr. Leonard's report is therefore at the same time enlightening and reassuring.

The report reveals numerous single, limited studies of important problems of English teaching. The chief value of these isolated studies probably lies in the techniques which they have developed rather than in their results. They should be checked by other teachers and the results studied comparatively. While it appears from the mere bibliography that certain of the problems have been adequately checked, examination of the studies themselves reveals that there is usually enough variation in the techniques to invalidate comparisons of results.

After analyzing the studies reported and his notes on his numerous conferences, Professor Leonard appends to his report a list of some of the most pressing and feasible research projects.

The Ninth Annual Congress of Psychology will be held at Yale University probably in August or September, 1929. This will be the first meeting of the Congress in America, previous meetings having been held at various intervals since 1889 in Paris, London, Munich, Rome, Geneva, Oxford, and Groningen. It is expected that the forthcoming Congress

will be truly international in character. Americans hope that the appointment of some foreigners for lecturers and lectureships can be arranged near the time of the Congress, so that foreign attendance can be increased and international solidarity furthered still more.

The following announcement recently appeared from the School of Education of the University of Southern California:

On the recommendation of the Faculty of the School of Education and endorsement of the Graduate Council, the Board of Trustees, at a meeting held June 7, took the following action:

"Upon the recommendation of the President, it was moved and unanimously carried that the School of Education, which has hitherto in its work for advanced degrees been regarded as a part of the Graduate School, be henceforth recognized as one of the separate professional schools of the University, operating on the same basis as the School of Law and the School of Medicine, and that the faculty be empowered to recommend for appropriate degrees those who complete satisfactorily the courses laid down for such recognition."

Under this action the Faculty of the School of Education was authorized to outline a three-year professional program leading to the degree of Doctor of Education. According to Dean Lester B. Rogers, of the School of Education of the University of Southern California, two factors have led to the establishment of the new degree. One was the desire to create a professional degree of equal weight and merit with the present Ph.D. degree, but with the emphasis on advanced professional preparation and mastery of educational principles and procedures rather than primarily on research. The other was to provide a master-teacher type of degree for the better preparation of teachers in subject matter fields for junior and other colleges. The new program will meet the needs of men and women in teaching positions who wish to continue their scientific study of Education, but who are not interested primarily in research.

Two curricula have been organized which lead to the degree of Doctor of Education. One is known as the Administrator type and the other will be known as the Master-Teacher type and will be primarily designed to prepare a new type of teacher on subject matter fields. The chief differences between the requirements for the new degree and the general university degree of Doctor of Philosophy will be in the emphasis placed on a broad and systematic knowledge of Education as a field of professional service, rather than on research. The thesis may be an interpretation or the organization and application of existing experience, as well as a contribution.

The reading knowledge of French and German, now required for the Ph.D. degree, will be changed to one modern language for the Master-Teacher curriculum, and may be waived entirely for the Administrator curriculum. In place of the language requirement the candidate will be expected to have full command of the tools that will be needed in his thesis work, including statistical methods, etc.

The School of Education will continue to give work leading to the Doctor of Philosophy degree under the general regulations of the Graduate Council. Hereafter, all candidates for credentials, candidates for the Master of Arts degree in Education, and candidates for the Doctor of Education degree will be admitted by way of the Admissions Committee, and will be under the entire direction of the Faculty of the School of Education.

The International Institute of Teachers College, Columbia University, and the Central Institute for Education and Instruction of Berlin announce a tour of Germany's educational institutions for the summer of 1929. The tour is organized by the Central Institute which is operated under the auspices of the Prussian Ministry of Education and the German Federal Ministry of the Interior. This is the first time that American teachers as a group have been afforded an opportunity, under official direction of

the German educational authorities, to gain a direct insight into the German school system. All types and levels of education will be observed, including allied activities such as the Youth movement and juvenile welfare.

The visitation of schools will begin on June 17, 1929, and will continue for six weeks, concluding just before the World Federation of Educational Associations' Conference in Geneva, which meets the last week in July. The tour will include visitation to the most prominent German cities, among which will be Bremen, Hamburg, Lübeck, Kiel, Berlin, Magdeburg, Leipzig, Dresden, Chemnitz, Gera, Jena, Weimar, Nürnberg, Munich, Stuttgart, Heidelberg, Frankfort, Weisbaden, Coblenz, Mainz, Bonn, Cologne, Düsseldorf, and Essen. The selection of members of the party will be in the hands of the International Institute of Teachers College, Columbia University, New York City. There will be twenty-five in the party and the membership will be limited to those who have some command of the German language. Members will assemble at Bremen or Hamburg, arriving by any route which they desire. Complete information may be had from Dr. Thomas Alexander, Teachers College.

The trend in superintendents' annual reports away from general discourse toward presentations of quantitative evidence on local educational conditions and progress is exemplified in the 1927-28 Report of the Public Schools of Burnham, Pennsylvania, prepared by Supervising Principal Elmer E. Sipe. Objective data, based upon the results of standardized tests, are presented analytically for arithmetic, reading, spelling, and handwriting. Factors causing retardation were also given careful study. Mental deficiency was found to rank seventh in a list of ten causes. Indifference and indolence on the part of pupils and lack of home support were found to be the chief factors in retardation. About one-tenth of the cases were due to physical defects.

An illuminating by-product of the handwriting test was the discovery that only ten pupils in the eighth grade could write the first stanza of *America* correctly. The report voices the suspicion that too much learning is of the parrot type, and recommends that *America* be included in the list of required poems. From a hundred or more garbled stanzas the following example is quoted in the report:

> Long may our fathers die
> With freedom's holy light,
> And from the mountain side
> Let freedom ring.

Under the auspices of the State Department of Public Instruction, committees of teachers are at work revising the various departmental courses of study for Indiana high schools. Several of the new courses are nearing completion.

An intensive three-year analysis of required teacher-training courses in Education at the University of Minnesota has just been completed by Dr. W. E. Peik of the College of Education. His first step was to analyze the various required courses into several hundred topics. He then secured from each instructor a time distribution on each topic. Finally, testimony on the value of each topic and each course was secured from graduates with teaching experience. A profile of each topic was then drawn, revealing the total time allotted to it, the amount of overlapping in the several courses, the degree to which the graduates recalled the topic, the value of the topic to them

in their teaching, and its value in their educational thinking.

Considerable overlapping was found in topics ranking both high and low in value to teachers in service. On the other hand, certain topics which were ranked high by graduates in the field were being virtually neglected in the required training courses. As a result of the investigation, two courses, formerly required, have been made elective and the important topics in them substituted for low-value units in other required courses. To secure the necessary total of credits in Education, teachers in training will elect from a list of courses including the two that were formerly required.

Faculty committees have been appointed to analyze and reorganize each required course in the light of Dr. Peik's profiles.

The coöperative study of the selection and organization of content and of methods of teaching which was under way all of last year in the public schools of Springfield, Missouri, will be continued during the coming year.

Children's interests and tastes in literature are constantly receiving both wider and more intensive treatment. A significant phase of the general problem, namely, the factors that affect children's interests in literature, has been pursued for the past two years by Dr. Grayson Kefauver, Assistant Professor of Education, University of Minnesota. Some of the factors studied by Dr. Kefauver were age, intelligence, grade, scholarship, home library (number of books and magazines), and language spoken by parents. Literature tests and graphic ratings were devised or adapted and the results analyzed and compared in various ways to isolate the critical factors in the reading tastes of pupils.

In the construction of scales for the evaluation of handwriting, composition, drawing, moral choices, religious attitudes, poems, and scores of other variables, investigators have used a unit of measure taken from the base line of the normal curve. This practice may be justified in terms of the Cattell-Fullerton theorem that equal differences are equally often noticed. Thurston[1] has shown that this is true only in case dispersions are equal. The problem has not infrequently arisen, "Do I have enough judges?" Probably "Get as many as you can" is one good answer. A minimum value to be taken into account in answering that question is here suggested as a working basis.

It is assumed that it is necessary to be practically certain that Sample A is really better than Sample B. This should be tested out, if samples are involved, at that interval in the scale at which there is most doubt. This will be the case in which the number of judges believing A better than B is smallest.

Let p represent the percent of judges believing A better than B.

Then, as Yule[2] has shown, the standard error (σ) of that measure may be represented by $\sqrt{\frac{p\,q}{n}}$ where n is the number of judges. If we are to be practically certain that the order of merit will hold, then $p - 3\sigma > 50$ or $p - 3\sqrt{\frac{p\,q}{n}} > 50$ which, solved for n, gives $n > \frac{9\,p\,q}{(p-50)^2}$

[1] Thurston, "Equally Often Noticed Differences," *Journal of Educational Psychology*, XVIII (May, 1927), 289.

[2] Yule, G. U. *An Introduction to the Theory of Statistics*, (Seventh Edition. London: Charles Griffin & Co., Ltd., 1924), p. 257.

TABLE I

Percent of Judges Believing Sample A better than Sample B; the Interval A-B being the Smallest in the Scale	55%	60%	65%	70%	75%	80%	85%	90%
Minimum Number of Judges for Practical Certainty	892	217	92	48	28	17	10	4

Table I has been prepared to indicate, in accord with this interpretation, the smallest number of judges which will be adequate for several values of p. It is indicated, for example, that the common practice of interpreting a scale unit in terms of 75 percent of the judges believing A superior to B, should not be based upon the choices of less than 28 judges. Of course, for small numbers of judges, the value of the standard error may well be called in question. Also, it must be borne in mind that the Cattell-Fullerton theorem does not work well at the extremes of the scale.

GOODWIN B. WATSON
Teachers College, Columbia University

REPORTING FAILING GRADES

Should college students who are doing inferior or failing work be notified of their delinquency? If they should be informed, how should the notification be made? Some college teachers are opposed to any effort to keep inferior students posted as to their academic standing. The argument presented is that a college student is far enough advanced in age and maturity to be able to know whether or not his work is satisfactory.

In the Central Missouri State Teachers College it has been the custom for some years to send written notices to students, in the middle of the term, who are doing inferior or failing work. It seemed necessary to make a study of this activity to determine whether it was advisable to continue such a procedure. It was thought that the opinions of the student body concerning this practice would be a good indication as to its effectiveness. Students were asked to give their opinions in writing concerning the notification of those doing a low grade of work. Care was taken to inform them that it would be unnecessary to sign their names to these statements. The students were further informed that no attempt would be made to check on any individual as to his opinions. They were urged to express their opinions freely. A total of 767 students were asked to express their opinions concerning whether some method of notification of those doing a low grade of work should be practiced. Table I shows the results of such a study.

TABLE I

DISTRIBUTION OF OPINIONS OF STUDENTS CONCERNING THE REPORTING OF INFERIOR OR FAILING GRADES

	FRESHMAN	SOPHOMORE	JUNIOR	SENIOR	TOTAL
Positively for Notifying	131	144	119	118	512
Positively Against Notifying	61	70	53	48	232
Both For and Against	0	7	5	11	23
Total Reporting	192	221	177	177	767

TABLE II

SHOWING OPINIONS OF STUDENTS AS TO METHODS OF NOTIFICATION

	FRESHMAN	SOPHOMORE	JUNIOR	SENIOR	TOTAL
Private Conferences with Insts.	74	68	31	35	208
Continue Present Plan	21	41	9	8	79
Send Notices in Plain Envelopes (Not use College Stationery)	3	7	6	4	20
Send out all Grades at Mid-term	4	1	2	2	9
Note on Mid-term Test Paper	9	2	1	1	13
				Grand Total—	329

Students were asked to make suggestions as to what they considered the best possible method of notifying them of inferior work. A few of the suggestions are listed in Table II.

The opinion of poor or failing students who have received notification, during the term, of inferior or failing grades should be a good index of the success of the plan. In a student body of 1908 students, 683 were notified of their poor work. They were informed on the slip sent them that the Dean of Women would be glad to advise with them. This conference was purely optional with the students. The Dean of Women desired to ascertain the cause of the low grade of work. Of this number 388 girls and 71 boys reported. Each student was asked his opinion concerning the notification. Approximately 90 percent of those reporting to the Dean of Women said they were glad to be notified of their standing during the term so that they could exert more effort on the subject in which they were deficient.

The practice in other colleges and universities should be another indication of the advisability of notifying inferior or failing students. A questionnaire was accordingly sent to 123 higher institutions. Table III shows the results of 109 replies. Fifty of them came from state teachers colleges, 25 from normal schools, 21 from universities, and 13 from academic colleges or parochial colleges. The distinction drawn here between teachers colleges and normal schools is that drawn by the United States Bureau of Education.

It is seen by Table III that 108 of the 109 schools deemed it advisable to notify students who are doing poor work; 104 make some practice of notification and five do not. The method of notification was varied. Forty-nine notified by mail; 75 made their notification by asking the students to hold conferences with their instructors; 42 made suggestions on examination papers and 22 asked the student to hold a conference with the

TABLE III

SHOWING DISTRIBUTION OF 109 COLLEGES AND UNIVERSITIES AS TO METHODS USED IN NOTIFYING STUDENTS DOING INFERIOR WORK

NOTIFICATION ADVISABLE		PRACTICE OF SCHOOLS		WHAT SYSTEM USED				
Yes	No	Yes	No	By Mail	Conference with Inst.	Sug. on Ex. Pap.	Conference with Dean	Other Methods
108	0	104	5	49	75	42	22	6 Post on Bulletin Bds. 4 Conf. with Committee 4 Notify Parents 4 Conf. with President

dean. Administrators sometimes showed disapproval if conference with the instructor was the only method employed in giving this information to students. The arguments presented were that not all professors would take a personal interest in the academic standing of their students and would fail to inform them of their low standing. The student would maintain at the end of the quarter or semester that he did not know of his low rating by his professor and supposed that he was doing satisfactory work.

It was the practice in some schools for the faculty to report both superior and inferior work on the assumption that the superior student is as much of a problem in the ordinary classroom instruction as the inferior one.

The faculty adviser was used in some schools in an attempt to get in closer touch with the student body. An adviser was appointed to each student when he entered school. The duty of this faculty member was to get acquainted, both socially and academically, with the students under his care and to remain their counselor and helper throughout the year. In some institutions this practice was used only in the Freshman class because the size of the student body rendered it impossible for the faculty to make acquaintance with such a large number of students.

The time of notifying students of their delinquency varied. The most common practice seemed to be to enlighten students of their standing at mid-term. A few were insistent that reports be made oftener than once a semester or quarter, while some declared that the report should be made at any time the delinquency was discovered.

It is evident that students are strongly in favor of some form of notification by the faculty if they are doing inferior work. Those who receive information of a low rating of their work seem deeply sensible of the kindness received and are urged by this act to exert themselves to secure a higher rank. The most prevalent opinion among the students was that a notification through a private conference with their professor would be of most value. It is also noted that the practice of this method ranks highest among the colleges and universities included in this study. The personal contact with professors was no doubt the contributing factor in favor of conferences with the instructor over the cold, formal method of notification by mail.

C. A. COLLINS
Central Missouri State Teachers College

The Research Division of the National Educational Association in coöperation with the Statistical Division of the Federal Bureau of Education, has published a large amount of practical and useful information. The Bulletins of the Research Division are always eagerly received. Vol. V, No. 4, on "The Advance of the American School System" is no exception. There is, however, in this last Bulletin on page 222 a table on the "Educational Rank of States by Two Methods" which must be very discouraging to the people of some states, provided they take the rankings seriously. Utah, for example, fell off in four years, from 1920 to 1924, from 5th place to 20th place, according to Phillips and from 6th place to 24th place according to Ayres. There is a good correlation between the Ayres and Phillips methods but it is a case of high correlation of errors; because we all know that if Utah was 5 or 6 from the top of the 49 states and the District of Columbia in 1920, it was not on the toboggan to the time of being

surpassed in four years by from 15 to 18 other states.

Idaho had an even swifter decline, going down from 2nd to 24th place by Phillips and from 5th to 29th place by Ayres. In short, about as many states moved down as there were states that moved up. From the nature of the ranking method this had to be; but improvement is delightful and easy to fall in with, so that the states with improved ranks may be satisfied, but those states with poorer ranks may wonder wherein they have so suddenly deteriorated. Ranking the states is a very questionable procedure. A better way is to show a measure of the states concerned with actual advancement rather than with the mere shifting of the order on a fixed scale.

The Index of Holding Power[1] is a less fluctuating measure of state, county, and large city school systems; and while the ranking of the states is not advocated, it may be operated with the Indices of Holding Power, just as with the Ayres and Phillips methods; but if the ranks are used there will be large fluctuations occasionally. This ranking has been done for 1922 and 1924. Also the states have been ranked according to the Average Number of Years Children Attend School (page 215 same bulletin).

CORRELATIONS OF RANKS OF STATES*

Ayres with Phillips 1924	r .926
Ayres with Phillips 1920	r .898
Ayres with Phillips 1910	r .939
Ayres 1910 with Ayres 1920	r .897
Phillips 1910 with Phillips 1920	r .875
Ayres 1920 with Ayres 1924	r .857
Phillips 1920 with Phillips 1924	r .898
Phillips 1924 with Av. No. Years children attend school	r .700
Index of Holding Power 1922 with 1924	r .939

* Pearsons' Product Moment Formula was used.

[1] "Index of Holding Power," *American Educational Digest*, August, 1927, and "Public School Progress of the States," *American School Board Journal*, October, 1927.

The standard deviation is exactly the same in all the distributions with the exception of the Indices of Holding Power, where, due to the same ranking of several states, there is a slight variation.

Educational ranks of the states has absolutely nothing to do with the subject of the Research Bulletin, Vol. V, No. 4, "The Advance of the American School System." By ranking the states, there is always a first and a forty-ninth, with all the other states somewhere between; but no advancement can possibly be indicated by the rank method. A shifting of the order of the states is all that the system reveals, and this is surely not "The Advancement of the American School System."

The Ayres Indices, revised method, which take into account the purchasing power of the dollar, will, if reliable, show improvement from year to year. The sum of the ranks of the ten Phillips components will show changes from year to year and appears to be more reliable than the ranks. The indices of Holding Power give a very reliable rating and indicates changes if there are any and moreover is simple and relatively free from error.

CORRELATIONS OF INDICES OR SUM OF RANKS (RAW SCORES)

Ayres with Phillips 1924	.428
Ayres 1920 with Ayres 1924	.329
Phillips 1920 with Phillips 1924	.95
Index of Holding Power 1922 with 1924	.933

Dr. Leonard P. Ayres gave us the first method of gross measurement of large school units and his work proved to be a great stimulant. Dr. Frank M. Phillips, chief of the Division of Statistics of the United States Bureau of Education, has greatly improved on the Ayres Method. Both methods are still fundamentally open to criticism, however, in that they are each made up of ten different components. When ten different things are combined it is difficult to know just what

the result of the combination really is. On the other hand, Holding Power is one concrete thing, but it is affected by everything that mars or enhances the schools. Holding Power has been used for many years and is now in use by many school superintendents, as evidenced by their reports. The Index of Holding Power has been suggested as a reliable measure of school progress. It is necessary to get the number of pupils under each grade of instruction. From these data, get the percent of pupils in each grade. The standard distribution of the percent of pupils in three groups of grades is as follows:

Pupils in the elementary grades 1-6.... 55%
Pupils in the middle grades 7-9..... 25%
Pupils in the senior grades 10-12..... 20%

The Index of Holding Power equals 100 minus the sum of the negative deviations from the standard, plus any positive deviation in the middle group.

If the data on distribution for eleven of the states, now largely estimated, by the Division of Statistics of the Bureau of Education, Department of the Interior, were correct, instead of estimated, as is the case for the other states, then the coefficient of correlation for the Indices of Holding Power for the 49 states and the District of Columbia for two successive years would undoubtedly be 98 or 99.

ERNEST C. WITHAM
Public Schools, Wilmington, Delaware

DICAL ROOM
RAL LIBRARY
Y. OF MICH.

NOV 9 1928

OCTOBER, 1928

Journal of Educational Research

CONTENTS

	Page
Measurement and Experimentation in Educational Methods — *V. A. C. Henmon*	185
The Retention of American History in the Junior High School — *Fowler D. Brooks and S. Janet Bassett*	195
The Vocabulary Load of Certain Secondary School Mathematics Textbooks — *H. H. Remmers and A. Grant*	203
A Study of a Modified Form of the Multiple Response Test — *Francis D. Curtis and Gerald G. Woods*	211
Unfamiliar Situations as a Difficulty in Solving Arithmetic Problems — *Carleton Washburne and Mabel Vogel Morphett*	220
A Table for Predicting the Validity and Reliability Coefficients of a Test When Lengthened — *Harold A. Edgerton and Herbert A. Toops*	225
A Comparison of Two Methods of Learning — *R. D. Russell*	235
An Experimental Investigation of the Relative Effectiveness of Two Plans of Supervised Study — *Harl R. Douglass*	239
Editorials	246
Reviews	249
Research Abstracts and Bibliographies	254
News and Communications	260

Copyright 1928, by Public School Publishing Co.

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879